University of Helsinki
Dissertationes Universitatis Helsingiensis 241/2025

Functional and Structural studies of Membrane-bound Pyrophosphatase and Its inhibitors

Jianing Liu

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Biological and Environmental Science of the University of Helsinki, for public examination in Biokeskus 2 Sali 2041, On 13th of June 2025, at 12:15 noon

Helsinki 2025

Pre-examiners

Docent Veli-Matti Leppänen, University of Turku

Professor Jens Preben Morth, Technical University of Denmark

Custos

Professor Adrian Goldman, University of Helsinki

Supervisor

Professor Adrian Goldman, University of Helsinki

Docent Henri Xhaard, University of Helsinki

Opponent

Professor Annalisa Pastore, Elettra Sincrotrone Trieste

Publisher: Helsingin yliopisto
Series: Dissertationes Universitatis Helsingiensis 241/2025

ISBN 978-952-84-1081-2 (Paper back)
ISBN 978-952-84-1080-5 (PDF)
ISSN 2954-2898 (Print)
ISSN 2954-2952 (PDF)
PunaMusta, Joensuu 2025

Abstract

Membrane-bound pyrophosphatases are homodimeric enzymes essential to the life cycles of bacteria, archaea, plants, and parasitic protists. They hydrolyze pyrophosphate into inorganic phosphate and pump H^+/Na^+ across membranes. Notably, mPPase are absent in humans, making them attractive therapeutic targets in pathogenic protozoan parasites responsible for severe diseases. To date, crystal structures of mPPases from *Vigna radiata* (VrPPase), *Thermotoga maritima* (TmPPase), and *Pyrobaculum aerophilum* (PaPPase) have been resolved in complex with various ligands to study their mechanisms and support structure-based drug design. The "pumping-before/after-hydrolysis" mechanism has been proposed, and several inhibitors have been identified. However, due to potential artifacts in crystal structures, the conformational changes of mPPase in solution remain unclear. Furthermore, limited knowledge about the pharmacological activity of inhibitors on mPPase in pathogenic parasites hinders progress in their application for drug design. In this thesis project, I employed biochemical and biophysical tools to study: (1) the conformational dynamics of well-characterized TmPPase with various bisphosphonates; (2) the expression, purification, and activity characterization of mPPase from *Plasmodium falciparum* (PfPPase); and (3) the identification of potential non-phosphorus inhibitors targeting PfPPase.

In Publications I and II, two crystal structures of TmPPase were solved in complex with bisphosphonates, providing evidence for catalytic asymmetry. Using double electron-electron resonance spectroscopy, I confirmed the conformational dynamics of mPPase in asymmetric states under various inhibitors and functionally relevant conditions. I proposed four distinct conformational ensembles of mPPase in the presence of different inhibitors. Combined with solid-supported membrane-based electrophysiology, these findings demonstrate that ion pumping requires pyrophosphate hydrolysis, supporting the "pumping-after-hydrolysis" model. These results further reinforce the concept of symmetry-breaking across the membrane. In Publications III and IV, I successfully expressed and purified PfPPase in insect cells. PfPPase was shown to exist in both dimeric and monomeric states, with enzymatic activity restricted to the dimeric form. I subsequently screened the inhibitory effects of fourteen antimalarial drugs, seven TmPPase inhibitors, and fourteen other compounds on PfPPase. Among these, three pyrazolo[1,5-*a*]pyrimidine-based TmPPase inhibitors retained micromolar IC_{50} values against PfPPase, while all other compounds were either inactive or showed minimal inhibition. To facilitate the structure-based drug development targeting mPPase from protozoan parasites, I fused BRIL, an engineered variant of apocytochrome b562a, to the C-terminus of PfPPase and formed a complex with an anti-BRIL antibody to enhance particle properties for cryo-electron microscopy. Initial structural data were obtained, laying the groundwork for future optimization.

Index

List of Original Publications

This thesis is based on the following publications:

I. **Liu, J**^., Shah, A^., Liu, X^., Wort, J^., Ma, Y^., Hardman, K., Johansson, N. G., Ribeiro, O., Brookfield, A., Bowen, A., Yli-Kauhaluoma, J., Xhaard, H., Jeuken, L. J. C., Goldman, A., Pliotas, C. & Vidilaseris, K. Conformational dynamics and asymmetry in multimodal inhibition of membrane-bound pyrophosphatases. *eLife* **13** (2024). DOI: 10.7554/elife.102288.1

II. Strauss, J., Wilkinson, C., Vidilaseris, K., de Castro Ribeiro, O. M., **Liu, J**., Hillier, J., Wichert, M., Malinen, A. M., Gehl, B., Jeuken, L. J., Pearson, A. R. & Goldman, A. Functional and structural asymmetry suggest a unifying principle for catalysis in membrane-bound pyrophosphatases. *EMBO Rep* **25**, 853–875 (2024). DOI: 10.1038/s44319-023-00037-x

III. **Liu, J**., Vidilaseris, K., Johansson, N. G., Ribeiro, O., Dreano, L., Yli-Kauhaluoma, J., Xhaard, H. & Goldman, A. Expression, purification and preliminary pharmacological characterization of the *Plasmodium falciparum* membrane-bound pyrophosphatase. *bioRxiv* 2025.03.30.646216 (*PLoS One* submitted)

IV. Johansson, N. G., Dreano, L., Vidilaseris, K., Khattab, A., **Liu, J**., Lasbleiz, A., Ribeiro, O., Kiriazis, A., Boije Af Gennäs, G., Meri, S., Goldman, A., Yli-Kauhaluoma, J. & Xhaard, H. Exploration of Pyrazolo[1,5-*a*]pyrimidines as Membrane-Bound Pyrophosphatase Inhibitors. *ChemMedChem* **16**, 3360–3367 (2021). DOI: 10.1002/cmdc.202100392

^Joint first author

The publications are referred to in the text by their roman numerals.

Description of key contribution by JL to the publication

I. **JL, KV, AG**, and **CP** developed the concept for this study. **JL** and **KV** performed atomic model building, structure validation, visualization, data interpretation, and wrote the manuscript. JL performed the electrometric measurements and analyzed the data. **AS, XL**, and **JLW** conducted the DEER experiments and analyzed the DEER data. **AG** and **CP** supervised the data analysis. All authors contributed to the review and editing of the manuscript

II. **JL** participated in sample preparation, performed the electrometric measurements, and contributed to the methodology and interpretation of the results presented in the manuscript.

III. **JL, KV, HX,** and **AG** developed the study concept. **JL** planned and optimized the experiments, prepared the samples, interpreted and visualized the data, and wrote the manuscript. **NGJ** provided the compounds, and **LD** validated the calculations. All authors contributed to the review and editing of the manuscript.

IV. **JL** participated in sample preparation, performed the inhibition assays, and reviewed the manuscript.

Abbreviation

ALE	Alendronate
ATC	2-amino-6-benzothiazolyl)methyl]-1H-indole-2-carboxamide
AWI	Air-water interface
BAK5	Affinity-matured sAB
BRIL	Apocytochrome b562
CHS	Cholesteryl hemisuccinate
CMC	Critical micelle concentration
CTF	Contrast transfer function
CryoEM	Cryo-electron microscopy
DDM	n-dodecyl-ß-D-maltoside
DEER	Electron-Electron Double Resonance Spectroscopy
DTT	Dithiothreitol
ETD	Editronate
GDN	Glyco-diosgenin
GFP	Green fluorescent protein
GPCRs	G protein-coupled receptors
IC_{50}	Half maximal inhibitory concentration
IDP	Imidodiphosphate
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LMNG	Lauryl maltose neopentyl glycol
MALs	Multi-angle light scattering
mPPase	Membrane-bound pyrophosphatase
	(1-Oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)
MTSSL	Methanethiosulfonate
NRD	Neridronate
PAM	Pamidronate
PDB	Protein data bank
PMSF	Phenylmethylsulfonyl fluoride
PPi	Pyrophosphate
RSD	Risedronate
RT	Room temperature
SDS	Sodium dodecyl sulfate

SEC	Size-exclusion chromatography
SMA	Styrene–maleic acid copolymer
SNR	Signal-to-noise ratio
SSM	Solid-supported membrane
TEV	Tobacco etch virus
ZLD	Zoledronate

1 Introduction

1.1 Pyrophosphatases

Pyrophosphatases (PPases) are essential enzymes that catalyze the hydrolysis of inorganic pyrophosphate (PP_i) into two orthophosphate (P_i) molecules. This reaction is critical for maintaining cellular energy homeostasis and facilitating biosynthetic pathways by ensuring the irreversible breakdown of PP_i, a byproduct first identified in various mammalian tissues in 1928. PP_i plays a vital role in numerous biochemical reactions and also acts as a regulatory molecule for various enzymes without directly participating in the reactions[1]. Excessive accumulation of PP_i is cytotoxic, disrupting the metabolic reactions that product PP_i, as well as DNA/RNA polymerisation and connective tissue matrix calcification[2-4]. Therefore, cell metabolism heavily depends on PPases activity across all kingdoms of life.

PPases are classified into three major protein families: Family I, Family II (sPPases), and membrane-integral pyrophosphatases (mPPases). sPPases are soluble enzymes with no evolutionary relationship, mPPases exhibit significant structural and functional differences from sPPases, despite catalyzing the same reaction[5].

1.1.1 Family I and II Pyrophosphatases

Family I PPases are extensively studied single-domain enzymes that are widely distributed across bacteria, archaea, and eukaryotes. These enzymes typically function as oligomers, most commonly adopting dimeric in eukaryotes or hexameric configurations in prokaryotes[6,7]. The first structure of a family I PPase structure was determined from *Saccharomyces cerevisiae* [8], revealing a dimeric organization and providing insights into the conserved active site conformation where PP_i binds. Their active sites require three to four divalent metal ions, such as Mg^{2+} or Mn^{2+}, to achieve catalytic activity[9].

Unlike the broadly distributed family I enzymes, family II PPases are primarily found in specific bacteria and archaea, suggesting they play specialized roles in the metabolic processes of these organisms[10]. The first structures of Family II PPases were solved from *Streptococcus mutans*, *Streptococcus gordonii* and *Bacillus subtilis*, they possess two domains of N-terminal and C-terminal domain per subunit, as opposed to the single-domain architecture observed in Family I enzymes[5,11,12]. Their active sites are located at the domain

interface and contain two conserved histidine residues, which confer a preferential requirement of Mn^{2+} and Co^{2+} over Mg^{2+} as metal cofactors[5,11,12]. In the presence of Mn^{2+} and Co^{2+}, these enzymes exhibited higher catalytic activity compared to Mg^{2+}, likely due to the superior ability of Mn^{2+} and Co^{2+} to accommodate the transition from five- to six-coordinate geometry during catalysis, a transition that Mg^{2+} handles less efficiently[5]. Despite this, the catalytic efficiency of family II PPases in the presence of Mg^{2+} is about 10-fold higher ($k_{cat:}$ ~2000 s^{-1}) compared to family I PPases ($k_{cat:}$ ~200 s^{-1})[5]. This enhanced activity is attributed to the tri-metal coordination of nucleophilic water in family II PPases, which facilitates more efficient deprotonation compared to bi-metal coordination observed in family I PPases[5]. Moreover, family II PPases employ a dissociative mechanism, wherein a strongly hydrated PO_3^- anion, such as metaphosphate, is formed prior to nucleophilic attack. In contrast, family I PPases lack an equivalent proton donor at the corresponding position and instead operate *via* an associative mechanism[5,12].

1.1.2 Integral Membrane-bound Pyrophosphatases

This thesis focuses on mPPases, a unique subclass of pyrophosphatases embedded within biological membranes. First discovered in the photosynthetic bacterium *Rhodospirillum rubrum* in 1966[13]. These enzymes are characterized by their structural complexity, with each subunit containing 15–17 transmembrane helices[14,15]. Unlike sPPases, which solely function to hydrolyze excess cytoplasmic PP_i, mPPases are primary ion pumps, coupling the hydrolysis of PP_i to the movement of Na^+ or/and H^+ across the membrane[5]. In general, ion pumping occurs from the cytosol, where PP_i hydrolysis takes place, to the periplasmic site when localized in the cell membrane, or to the organelle lumen when embedded in organelle membranes. In terms of hydrolysis, mPPases exhibit significantly slower activity, with a k_{cat} of approximately 10 s^{-1}, compared to sPPases[5] This reduced catalytic efficiency is primarily attributed to the superior nucleophile activation in sPPases, where the metal-coordinated water molecule is more effectively activated than the aspartate-coordinated water in mPPases[5]. Moreover, the capture of PP_i within the metal coordination cage in mPPases is less efficient than in sPPases, where the leaving group is primarily stabilized through interactions with side chains[5,13].

mPPases can be classified based on their ion selectivity (H^+, Na^+ and dual Na^+, H^+) and cation dependence[16-18] (**Table 1**). They are widely distributed across all domains of life, with the exception of multicellular animals [5]. Furthermore, they are considered among the earliest enzymes to couple phosphoanhydride bond formation and hydrolysis to the generation of chemiosmotic potential[19]. Phylogenetic analyses suggest that the utilization of PP_i as a chemical energy source predates the evolutionary adoption of ATP as the primary energy currency[19]. Sequence analysis further suggests the first mPPases were Na^+ mPPases and H^+ mPPases evolved four times from Na^+ mPPases, and the evolution of dual Na^+/H^+ mPPases is likely separate to the evolution of H^+ mPPases[18,20]. In terms of cation

dependence, K^+-independent mPPases have only been identified in H^+ mPPases[21]. All other mPPases require K^+ for maximal catalytic activity. In the following discussion, to facilitate the comparison of residues across mPPases, I will use the Ballesteros and Weinstein numbering system[22]. This system ensures that all functional residues across different organisms share the same index. Denoted as $X^{Y.Z}$, it references the amino acid (X), the helix number (Y), and the offset of well-conserved residue in the centre of the helix (at position 50 as per definition) (Z).

Table 1. Subfamily classification of mPPase

Ion selectivity	Monovalent dependence	Semi-conserved glutamate location	Cationic centre occupation	Regulation	Example
Na^+	Na^+, K^+	$E^{6.53}$	K^+ ($A^{12.46}$)		*Thermotoga Maritima* (TmPPase)
H^+	K^+	$E^{6.57}$	K^+ ($A^{12.46}$)		*Vigna radiata* (VrPPase)
	-	$E^{6.53}$	$K^{12.46}$		*Pyrobaculum aerophilum* (PaPPase)
Dual Na^+ and H^+	Na^+, K^+	$E^{6.53}$	K^+ ($A^{12.46}$)	Na^+	*Bacteroides vulgatus* (BvPPase)
	Na^+, K^+	$E^{6.53}$	K^+ ($A^{12.46}$)		*Clostridium leptum* (ClPPase)

Overall, Na^+ and dual Na^+/H^+ mPPases are prokaryotic mPPases[23,24]. In contrast, H^+ mPPases are also present in eukaryotes, including plants and parasitic protozoa[23,24].

1.2 Function of mPPases in prokaryotes and eukaryotes

In prokaryotes, mPPases are localized in cell membranes, where they function as critical energy-conserving devices, particularly under low-energy conditions such as cold or drought stress[25]. Consequently, they are predominantly found in obligate anaerobes like *Bacteroides* and deep-sea organisms (*Thermotoga maritima* and *Pyrobaculum aerophilum*), which thrive in harsh environments and grow optimally at temperatures ranging from 70 - 90°C[26-28]. Consistent with this, the melting temperature of TmPPase is approximately 70°C[29]. The remarkable stability of TmPPase makes it well-suited for various structural and functional studies, as explored in this thesis.

In plants, mPPases are predominantly localized in the tonoplast of the vacuole[23,24]. H^+ mPPases are abundant and play critical roles in plant maturation and ion homeostasis, functioning in tandem with vacuolar ATPases (V-ATPases), which generate the required

membrane potential. However, they do not exhibit a complementary relationship; knockout of V-ATPases disrupts normal cellular functions despite the compensatory increase in mPPase activity[30]. This is likely because mPPases function as a backup system under extreme stress conditions and are insufficient to fully compensate for the critical roles of V-ATPases under normal physiological conditions[31]. Overexpression of mPPases in *Arabidopsis thaliana* has been shown to enhance cell division and promote plant growth[32,33]. Conversely, knockout mutants of mPPase exhibit impaired root and shoot development due to reduced auxin trafficking driven by disrupted H^+ gradients[34]. Furthermore, the overexpression of mPPases has been shown to enhance drought and salinity tolerance by increasing vacuolar membrane potential, which facilitates osmotic potential and water uptake[35-39].

The life cycle of protozoan parasites heavily depends on mPPases for successful transitions between vector and host environments, as well as for adaptation to diverse host systems[40]. Specifically, in acidocalcisomes—small, acidic storage organelles where mPPases play a critical role—the internal environment acts as a reservoir for cations such as Ca^{2+}, Na^+, and K^+, as well as polyphosphates (Poly-P)[41,42]. The regulation of H^+ across the membrane is essential for maintaining osmotic homeostasis and acidifying the acidocalcisome[43,44], as its low pH is crucial for proper functionality. A knockout study of H^+-mPPase in *Trypanosoma bruceii* revealed that the loss of acidity leads to a tenfold reduction in stored Poly-P levels, resulting in detrimental osmotic changes[44]. Moreover, disruption of H^+ pumping — either by preventing its removal from the cytosol or by impairing its release from the acidocalcisome under acidic and alkaline conditions, respectively, adversely affected cellular pH regulation, cell growth rate, and final cell density. Beyond its role in acidocalcisome acidification, studies on H^+-mPPase knockdown and knockout strains of *Toxoplasma gondii* and *Plasmodium falciparum* suggest that mPPases also mediate *in vitro* asexual blood-stage growth and contribute to parasitic virulence[45,46].

1.3 Structural basis of mPPase

1.3.1 Structural overview

The crystal structures of VrPPase and TmPPase were first determined in 2012[15,47]. This breakthrough paved the way for the determination of various mPPase structures bound to ligands or ions (**Table 2**). However, until recently, these studies were limited to just these two organisms. In 2023, the structure of PaPPase in complex with imidodiphosphate (IDP) was solved [48], making an advancement in the field.

Table 2. Overview of mPPase structures

Species	Name	Conformation	Ligand/Ions	Resolution (Å)	PDB
Thermotoga Maritima	TmPPase:IDP	Closed IDP-bound	IDP, 5 Mg^{2+}, Na^{+}	3.5	5LZQ
	TmPPase:WO_4	relaxed product-bound	WO_4^{2}-, 2 Mg^{2+}	4.0	5LZR
	TmPPase:Ca	Resting	Ca^{2+}, Mg^{2+}	2.6	4AV3
	TmPPase:2P_i	Product-bound	2 P_i, 4 Mg^{2+}	4.0	4AV6
	Time-resolved TmPPase (t = 0 – 3600 s)	Resting (0-60 s), substrate-bound (300 s, 600 s), substrate and product-bound (3600s)	300 and 600 s: PP_i, 4 Mg^{2+}; 3600 s: P_i, PP_i, 5 Mg^{2+}	2.7, 2.5, 4.0, 3.8, 4.5	8B21-8B24
Vigna radiata	VrPPase:IDP	Closed IDP-bound	5 Mg^{2+}, K^{+}	2.4	4A01
	VrPPase:P_i	Relaxed product-bound	P_i, 2 Mg^{2+}	3.5	5GPJ
	VrPPase:2P_i	Product-bound	2P_i, 5 Mg^{2+}	2.3	WT (6AFS), $E^{6.57}Q$ (6AFT), $L^{12.64}M/K$ (6AFU, 6AFV), $T^{5.35}D$ (6AFW), $E^{5.33}A/S/H$ (6AFX-Z)
Pyrobaculum aerophilum	PaPPase:IDP	Closed IDP-bound	IDP, 5 Mg^{2+}	3.8	8B37

Overall, the solved mPPase structures exhibit similarity due to their high sequence identity (23-56%) and their function as homodimers[24], with each monomer comprising 16 transmembrane helices (TMHs). These helices are arranged in two concentric layers: an inner ring formed by TMHs 5-6, 11-12, and 15-16, and an outer ring consisting of TMHs 1-4, 7-10, and 13-14. Each monomer contains four functional regions: the active site, the coupling funnel, the ion gate, and the exit channel.

The active site of mPPases is a large hydrolytic centre extending into the cytoplasm (**Figure 1**). Despite variations among different mPPase subfamilies (**Table 2**), the active site contains highly conserved aspartate, asparagine and lysine residues ($K^{5.58}$, $D^{5.61}$, $D^{5.65}$, $D^{6.35}$, $D^{6.39}$, $D^{6.43}$, $D^{11.57}$, $D^{12.39}$, $N^{12.43}$, $D^{15.61}$, $K^{15.65}$, $D^{16.32}$, $D^{16.35}$, $K^{16.38}$ and $D^{16.39}$),which are essential for PP_i binding and hydrolysis (**Figure 1A-B**). The overall structure of conserved residues in the IDP-bound active site of Tm/Vr/Pa-mPPases is nearly identical (**Table 3**). Among these residues, aspartate residues interact indirectly with PP_i by coordinating with

five Mg^{2+} ions, forming a metal cage that captures PP_i. In contrast, lysine residues interact directly with PP_i, stabilizing it within the binding site. Mutagenesis of these active site residues typically results in a loss of enzymatic activity [15,49-52].

Table 3 Alignments of key residues between chains of different mPPase structures

Chain	$5LZQ_A$:$4A01_A$	$5LZQ_A$:$8B37_A$	$4A01_A$:$8B37_A$
C_α RMSD (Å)	0.37	0.81	0.73
The superposition of residues (($K^{5.58}$, $D^{5.61}$, $D^{5.65}$, $D^{6.35}$, $D^{6.39}$, $D^{6.43}$, $D^{11.57}$, $D^{12.39}$, $N^{12.43}$, $D^{15.61}$, $K^{15.65}$, $D^{16.32}$, $D^{16.35}$, $K^{16.38}$ and $D^{16.39}$)			

The coupling funnel is located between the active site and the ion gate and is lined with highly conserved charged residues ($R^{5.50}$, $D^{6.43}$, $D^{6.50}$, $D^{11.50}$, $K^{12.50}$, $K^{16.38}$ and $D^{16.39}$) (Figure **1C**). These residues form an ionic network that facilitates ion translocation through the hydrophobic protein region into the ion gate, coupled with PP_i hydrolysis[15,47]. Mutations in these residues, such as $D^{6.50}$ or $I^{12.54}$, result in reduced enzymatic activity and a complete loss of ion transport[52].

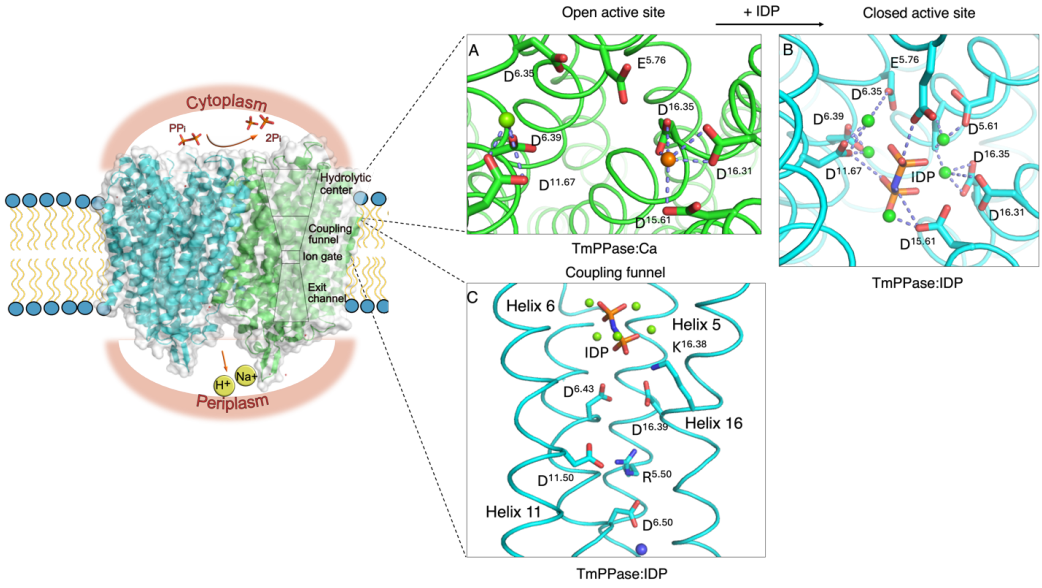


Figure 1. Overview of mPPase functional core regions with a focus on the active site and coupling funnel. (A) Close-up view of the open active site of TmPPase (green): Residues in the active site with Mg^{2+} and Ca^{2+} coordinated (dashed lines). Ca^{2+} (orange sphere) and Mg^{2+} (green spheres). (B) Close-up view of the closed active site of TmPPase (cyan): Residues in the active site with IDP coordinated (dashed lines) in a Mg^{2+} metal cage (green spheres). (C) Close-up view of the coupling funnel of TmPPase (cyan), extending from active site to Na^+ ion gate. Key residues are displayed along the translocation pathway. Na^+ (blue sphere) and Mg^{2+} (green spheres) are shown.

The ion gate acts as an Na^+/H^+ selectivity filter before ion pumping[14]. This functional core comprises charged and hydrophilic residues ($S^{5.43}$, $D^{6.50}$, $E^{6.53}$, $S^{6.54}$, $E^{6.57}$, $D/N^{16.46}$ and $K^{16.50}$) (**Figure 2A-C**). Of these, all residues except $S^{5.43}$ are directly involved in forming

the Na^+/H^+ binding site. However, ion selectivity primarily depends on the interaction between $S^{5.43}$ on helix 5 and the semi-conserved glutamate ($E^{6.53/57}$), as proposed in our recent work[48]. Briefly, in the structure of H^+-PaPPase, helix 5 adopts a slightly straighter conformation (**Figure 2D**), creating additional space for the protonated $E^{6.53}$ to interact with $S^{5.43}$. In Na^+-TmPPase, $S^{5.43}$ is pushed away from $E^{6.53}$ to prevent steric clashes, thereby forming an Na^+ binding site. In H^+-VrPPase, the downward shift of $E^{6.53}$ to $E^{6.57}$ generates sufficient space for $S^{5.43}$ to interact with the protonated $E^{6.53}$.

The exit channel, which connects to the ion gate, facilitates ion release but is not highly conserved in sequence among mPPases (**Figure 2E**). The only well-conserved residue in this region is $L^{12.64}$, which forms a hydrophobic gate with $V^{16.53}$, $S^{6.61}$ and $L^{5.40}$, functioning as a checkpoint to prevent ion backflow (**Figure 2E**) [15]. Mutations of this residue, such as $L^{12.64}$A/M/K in VrPPase, result in narrower exit channels, leading to reduced ion pumping and hydrolysis activity[53]. The open state of the exit channel is thought to represent the transient moment of ion release but has not yet been stabilized for structural determination. To date, the widest gate-open state has been observed in the VrPPase*:*$2P_i$ structure, which mimics the PP_i hydrolyzed state, featuring an approximately 4 Å pore size[53]. In this thesis, conformational changes in the exit channel were studied by measuring the inter-spin distance of residues within the exit channel using double electron-electron resonance (DEER) spectroscopy.

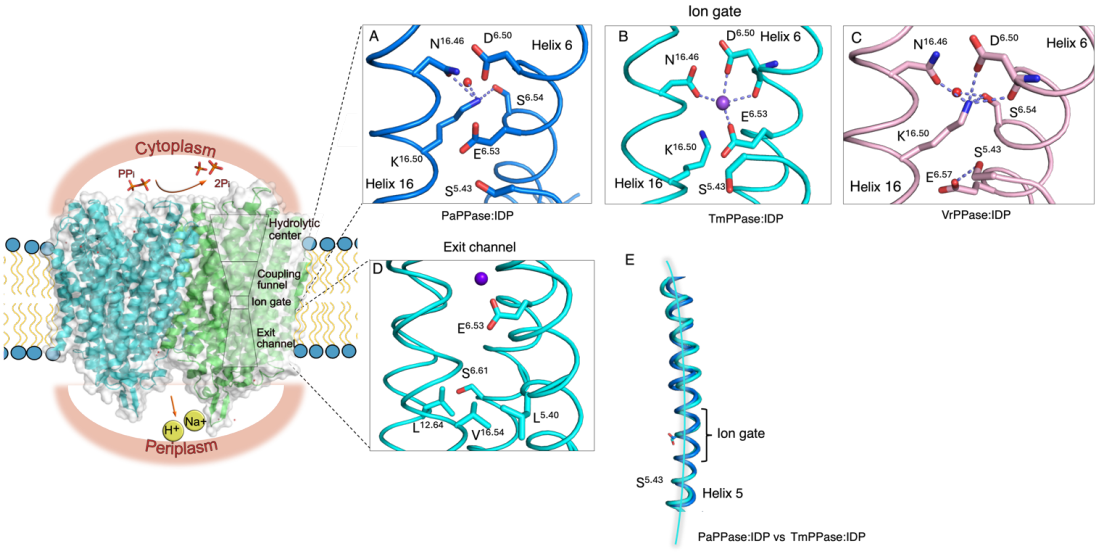


Figure 2. Overview of mPPase functional core regions with a focus on the ion gate and exit channel. (A) Close-up view of the ion gate in H^+-PaPPase:IDP structure (blue) shows the residue configuration when the active site is fully closed. Dashed lines indicate key interactions. (B) Close-up view of the ion gate in Na^+-TmPPase:IDP structure (cyan) shows the residue configuration when the active site is fully closed. Dashed lines indicate key interactions. Na^+ shown as purple sphere. (C) Close-up view of the ion gate in H^+-VrPPase:IDP structure (pink) shows the residue configuration when the active site is fully closed. Dashed lines indicate key interactions. (D) Close-up view of the hydrophobic gate at the top of the exit channel in

TmPPase:IDP structure. (E) Comparison of helix 5 orientation in PaPPase:IDP to TmPPase:IDP PaPPase:IDP (blue) and TmPPase:IDP (cyan)

1.3.2 Inter-subunit communication

It is now understood that the functional core regions of mPPases are all located within a single monomer. Early radiation inactivation experiments conducted in the 1990s revealed that the functional unit of mPPase operates as a dimer, where an impaired subunit can detrimentally affect the functionality of the otherwise unaffected subunit[54-56]. Dimerization is therefore believed to play a crucial role in ion pumping and PP_i hydrolysis. This is further supported by mutations in various residues at the dimer interface, which either inactivate the protein or result in "loose-coupling" mutants[51]. Meanwhile, kinetic studies have revealed that the substrate-binding affinities at two active sites are not identical[48,57,58] (**Table. 3**). Typically, PP_i binding to the first active site affects binding to the second active site ($K_{m,1} < K_{m,2}$) and reduces the hydrolysis velocity at the other site. The significantly lower V_1 in PaPPase compared to TmPPase and BvPPase may be attributed to its subclass affiliation (K^+-independent PaPPase vs. K^+-dependent TmPPase and BvPPase) or differences in membrane environments (detergent-solubilized PaPPase vs lipid-based TmPPase and BvPPase)[48,57,58]. The V_2 values for TmPPase and PaPPase are zero and approximately equal to V_1, respectively. This result is attributed to the fact that the steady-state kinetic experiments for these thermophilic enzymes were conducted at 40°C, a temperature where they remain reasonably active, despite their optimal temperature exceeding 70°C[48,57]. However, the instability and reduced solubility of Mg_2PP_i at elevated temperatures introduce significant error in determining V_2, as the solubility of Mg_2PP_i at 40°C is 800 μM lower than the K_{m2} of TmPPase and PaPPase[48,57]. Overall, these measurements indicate that in mPPase, only one active site predominantly functions at a time, rather than both sites operating simultaneously.

Table 4 Kinetic parameters for PP_i hydrolysis for three different mPPases

Parameters	TmPPase (Na^+) [57]	BvPPase (Na^+ and H^+) [58]	PaPPase (H^+) [48]
Organism	*Thermotoga maritima*	*Bacteroides vulgatus*	*Pyrobaculum aerophilum*
K_{m1} (μM)	1.41 ± 0.04	23 ± 2	2.42 ± 0.21
K_{m2} (μM)	1400 ± 100	800 ± 4	449.8 ± 828.4
V_1 (μmol $PP_i \cdot mg^{-1} \cdot min^{-1}$)	0.4 ± 0.005	0.28 ± 0.01	0.013 ± 0.0003
V_2 (μmol $PP_i \cdot mg^{-1} \cdot min^{-1}$)	~ 0	0.060 ± 0.04	0.009 ± 0.003

Despite evidence suggesting that mPPases function as a dimer with functional asymmetry, most crystal structures appear symmetric. This symmetry may be attributed to crystallization conditions, particularly the use of high concentrations of inhibitors (*e.g.*, 4 mM IDP), which can occupy both active sites, potentially leading to artefactual symmetry[14,15]. The first asymmetric structure of TmPPase was obtained in the presence of

N-[(2-amino-6-benzothiazolyl)methyl]-1H-indole-2-carboxamide (ATC)[59]. A dimer of ATC (ATC_2) selectively binds to one of the two monomers (*e.g.*, monomer A), stabilizing a locked conformation of the exit channel through interactions with loop6-7, loop8-9, and loop12-13 (**Figure 3A**). This binding event is facilitated by the downward motion of TMH 12, which induces the reorientation of loop12-13, a conformational change that occurs exclusively when PP_i or its analogue is bound at the active site. The subsequent upward displacement of TMH 13 at the dimer interface triggers a coordinated movement of TMH 13 in monomer B, enhancing the affinity of active site B for substrate binding. However, full catalytic activity is impeded due to the locked conformation imposed by ATC_2. The corresponding binding region of ATC_2 in monomer B is disordered (**Figure 3B**). RMSD/Cα values differ by up to 1.6 Å for certain loops at the exit channel, while the active sites remain structurally identical. Overall, this structure provides the first structural evidence that the mPPase dimer operates as an integrated machine through intra-subunit communication rather than as two independent units.

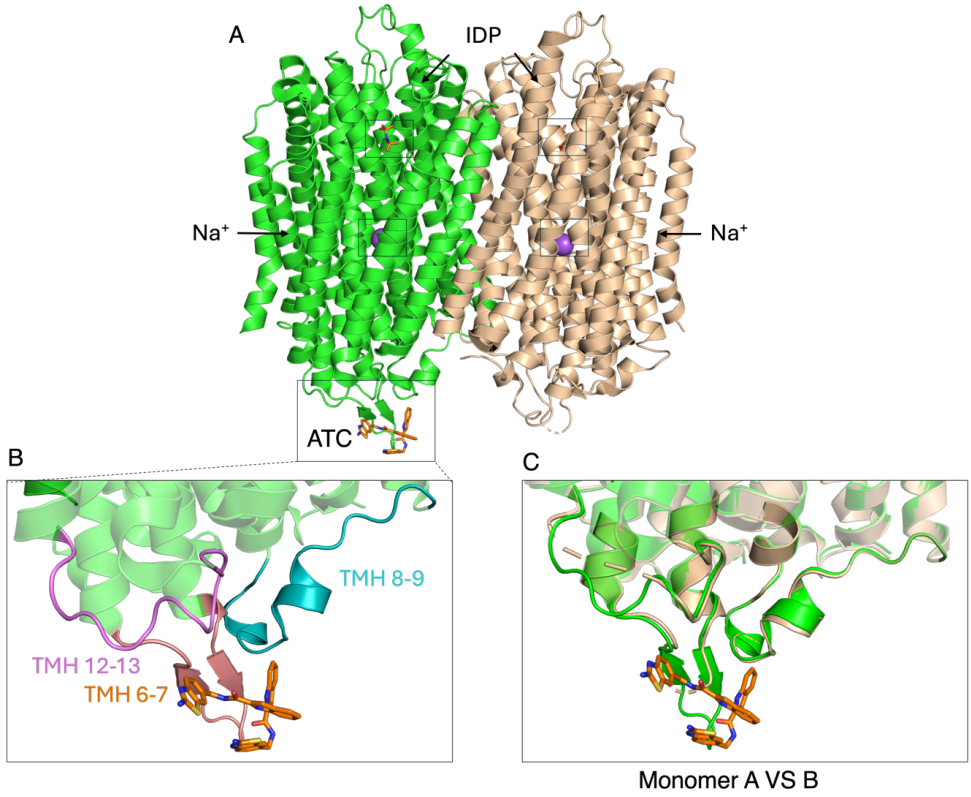


Figure 3. The asymmetric structure of TmPPase:ATC. (A) ATC dimer interaction with TMH6-7 (orange), TMH8–9 (cyan), and TMH12–13 (pink). (B) Comparison of the ATC binding site in monomer A (green) with the corresponding region in monomer B (orange).

The first asymmetric structure of mPPases at the active site was solved by Strauss and co-workers using time-resolved crystallography (**Figure 4**)[48]. TmPPase was incubated with PP_i

and cryo-trapped, effectively slowing the reaction to capture the hydrolysis steps over a period of 3600 seconds. At 60 seconds, no significant changes were observed in monomer A and B (**Figure 4A**), and the 0-60s structure remains symmetrical in the resting state, identical to the TmPPase:Ca structure (PDB: 4AV3). By 300 seconds, a positive mFo-dFc density appeared in the active site of monomer A, best fitting to PP_i; however, the PP_i had not yet reached its final binding pose, and the overall structure remains very similar to the TmPPase: Ca structure. After 600 seconds, PP_i reached the canonical position for hydrolysis. Concurrently, conformational changes were observed in the active site, including an inward movement of loop5-6 and a downward shift of TMH 12 (**Figure 4C**). Monomer B remained in the resting state until 3600 seconds (**Figure 4B**). Although the resolution of the 3600s-structure is poor, monomer A appears to still contain PP_i, while the density in monomer B is best fitted to $2P_i$ (**Figure. 4B)**. In summary, the time-resolved crystal structures provide snapshots of different states during the catalytic process, consistent with functional studies. In this thesis, further studies were conducted to further support the catalytic asymmetry under physiological conditions.

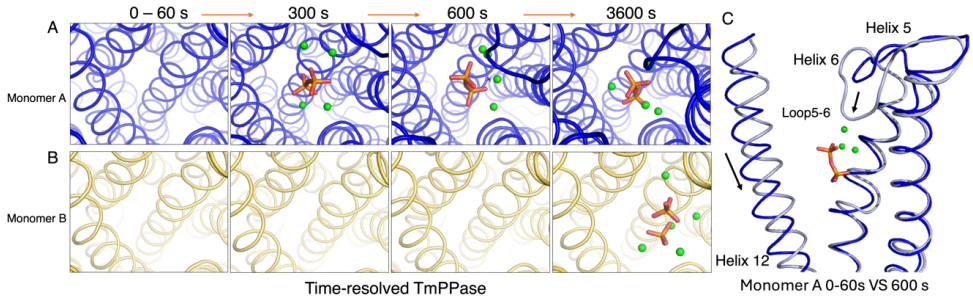


Figure 4. Time-resolved TmPPase structures in active site at different time-points. (A) TmPPase active site of monomer A at different time-points (dark blue); PP_i shown as a stick model and Mg^{2+} shown as green spheres. (B) TmPPase active site of monomer B at different time-points (yellow); $2P_i$ shown as a stick model and Mg^{2+} shown as green spheres. (C) Helix 12 reorientation and loop5-6 closure upon substrate binding in the 0 – 60 s and 600 s structure. TmPPase (0 – 60 s) shown in dark blue and TmPPase (600 s) shown in light blue. Black arrows highlight major structural changes.

1.4 Mechanism of Pumping and Hydrolysis

The mechanism of mPPase activity has been refined multiple times with increasing evidence from functional and structural studies[15,20,47]. Currently, researchers concur on a mechanism involving half-of-the-sites reactivity, intra-subunit communication, and exit channel motions as discussed above. Until recently, the primary debate centres on the sequence of hydrolysis and ion pumping during the catalytic cycle[14,60]. Two opposing mechanisms have been proposed: "pumping-after-hydrolysis" and "pumping-before-hydrolysis" (**Figure 5**).

The "pumping-after-hydrolysis" model, also known as "direct coupling", was proposed by Baykov *et al.*[60] (**Figure 5A**). In this model, the released H^+ originates from the nucleophilic

water molecule after it attacks PP_i. For Na^+ pumping, it is suggested that gate-bound Na^+ is propelled into the exit channel by H^+ *via* a 'billiard-type' mechanism[60,61]. In contrast, the "pumping-before-hydrolysis" model, also referred to as the "binding change" model, was first proposed by our group[14]. This model posits that H^+ and Na^+ pumping occur independently of hydrolysis; ions may originate from the medium or preceding hydrolysis events. However, active-site closure in the presence of PP_i is required to drive the cation out of the ion gate[14]. The increased negative charges along the coupling channel triggers the downward movement of TMH 12, followed by the deprotonation of $D^{6.43}$ and $D^{16.39}$, which activates the water nucleophile and ultimately leads to hydrolysis[14].

Support for ion pumping preceding hydrolysis comes from electrometric studies on H^+-VrPPase[14]. Li and co-workers measured proton current signals across membranes containing VrPPase upon the addition of PP_i and its analogues on gold-coated sensors[14]. In addition to PP_i, a tenfold lower current was detected in the presence of IDP. Based on these findings, they proposed that the ion pumping can occur independently of hydrolysis but requires the binding of PP_i or structurally similar analogues[14]. However, the interpretation lacks sufficient robustness to definitively rule out the "pumping-after-hydrolysis" model. The lower signal induced by IDP could result from partial ion transport across the membrane or be an artifact of charged amino acid reorientation within functional regions[60]. In recent studies, Baykov and coworkers used a stopped-flow pyranine assay to directly monitor proton transport across vesicles containing mPPase from *Desulfitobacterium hafniense*[62,63]. By measuring fluorescence changes of the pH-sensitive dye pyranine, they confirm that proton pumping occurs exclusively in the presence of PP_i, rather than its analogues[62,63]. Accordingly, in this thesis, electrometric studies were conducted on Na^+-TmPPase to further validate "pumping-after-hydrolysis" model.

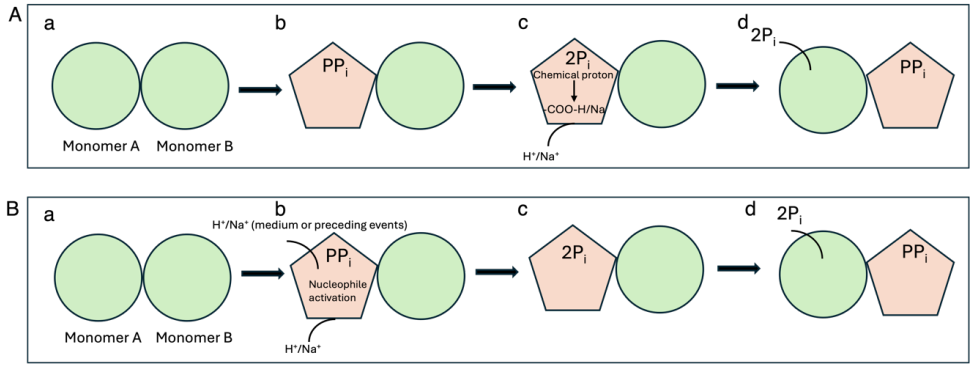


Figure 5. Schematic mechanisms of pumping and hydrolysis. (A) "pumping-after-hydrolysis" mechanism. Both monomers are in the resting state (a). With the addition of PP_i, one monomer binds PP_i and undergoes a conformational change to a closed state, while the other remains open (b). Next, the nucleophilic attack of a water molecule on PP_i releases a chemical proton, which electrostatically displaces the H^+/Na^+ ion bound to the Glu residue. The H^+/Na^+ ion then translocate into the exit channel (c); Finally, the Pi product is released from one monomer, and the other one became capable of initiating its own activity upon binding PP_i (d). (B) "pumping-before-hydrolysis" mechanism. Both monomers are in the resting state (a). H^+/Na^+ from medium

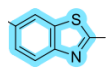
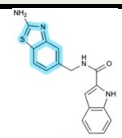
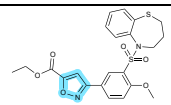
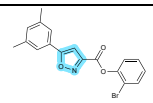
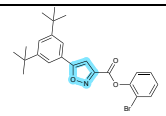
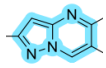
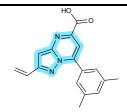
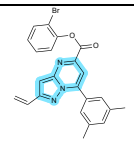
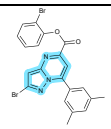
or preceding hydrolysis events rapidly crosse one monomer of mPPase, and the negative charges generated by ion pumping trigger the hydrolysis of PP_i in the same monomer (b). The Pi product binds to one monomer (c). Finally, Pi product is released, enabling the other one monomer became catalytically competent upon PP_i binding (d). The active site status is defined by shape.

1.5 MPPase inhibitors against protozoan parasites

Protozoan parasites, such as *Plasmodium spp.*, *Leishmania spp.*, and *Trypanosoma spp.*, present significant global health challenges. Climate change is accelerating the spread of insect vectors into new regions, thereby increasing the risk of severe, life-threatening diseases in the coming years[64]. In 2022, an estimated 249 million malaria cases, 0.7–1 million cases of leishmaniasis, and 6–7 million cases of Chagas disease were reported[65]. Moreover, current treatments are often costly and associated with severe side effects, limiting their accessibility in low-resource settings. The combination of drug resistance, limited treatment options, and expanding disease transmission underscores the urgent need for novel, affordable, and effective therapeutic strategies to combat these devastating diseases.

As discussed above, mPPases are found in eukaryotic protozoan parasites (see 1.2), but are absent from multicellular animals, making them attractive drug targets. To date, no approved drugs specifically targeting mPPase function in parasites are available. Our research focuses on identifying non-phosphorus-based inhibitors, as non-hydrolysable PP_i analogues can be recognized by human inorganic pyrophosphatases, which also utilize PP_i[66]. Our collaborators, Johansson and co-workers, identified three scaffolds from a proprietary library: benzothiazoles, isoxazoles, and pyrazolo[1,5-*a*] pyrimidines[59,67,68]. These scaffolds, which contain heterocycles with nitrogen, oxygen, and/or sulphur in five- or six-membered rings (**Table 4**), served as the initial fragment screening hits for synthesizing several compounds. The synthesized compounds were subsequently evaluated for their inhibitory effects on TmPPase by measuring its hydrolytic activity to determine IC_{50} values. The most potent compound identified so far is ATC, a benzothiazole derivative, with an IC_{50} of 1.7 μM[59]. However, ATC shows no activity against PfPPase[59]. Based on homology modelling, this lack of efficacy is attributed to structural differences in the binding region, particularly in loop6-7 and loop12-13, where ATC_2 is positioned[59]. Notably, loop6-7, which contributes two-thirds of the binding interface, is significantly shorter in PfPPase, impairing the effective binding of ATC. Apart from ATC, three isoxazole-based and three pyrazolo[1,5-*a*] pyrimidine-based inhibitors exhibit IC_{50} values below 10 μM[67,68]. In this thesis, part of the research was dedicated to assessing the inhibitory activity of various scaffold compounds against PfPPase.

Table 5. The inhibitory activity of representative compounds against TmPPase

Scaffold	Structure	IC_{50} (TmPPase,μM)	Reference
 Benzothiazoles		1.7	59
 Isoxazoles		5.4 [5.1–5.7]	67
		6.7 [6.6–6.8]	
		10 [7.3–13]	
 Pyrazolo[1,5- *a*] pyrimidines		14 [13–15]	68
		18 [17–19]	
		14 [13–15]	

1.6 The approach to structural and conformational investigation

Understanding protein mechanism requires integrating static structures with dynamic conformations, necessitating the use of diverse biochemical and biophysical tools. The rapid advancement of software has further propelled these studies. In this thesis, I provide a brief overview of four key approaches employed in the structural studies.

1.6.1 1Double electron-electron resonance spectroscopy

Double electron-electron resonance (DEER), also known as pulsed electron double resonance (PELDOR) spectroscopy, is a powerful technique for measuring precise inter-spin distances between spin-labelled sites in biomolecules[69-71]. DEER is highly effective for studying the structure, dynamics, and conformational states of integral membrane proteins, such as ion channels, transporters, outer membrane proteins, and receptors, within their native environments[72-78]. It serves as a valuable complement to structural biology techniques, especially for studying flexible, multi-state, or membrane-embedded systems that cannot be fully captured by static methods[79-82].

Paramagnetic spin labels (*e.g.,* MTSSL or other nitroxide probes) are introduced at specific cysteine residues within the protein using site-directed mutagenesis (**Figure 6A-B**)[83,84]. Continuous wave electron paramagnetic resonance (CW EPR) spectroscopy is then used to confirm the successful incorporation of spin labels and assess their mobility (**Figure 6C**)[85,86]. In the subsequent PELDOR/DEER experiment, the sample is recorded to optimize the frequencies of the pump and detection pulses (**Figure 6D**)[87]. By applying two distinct microwave frequencies, specific subpopulations of spin centres within the sample are selectively excited, allowing them to function as either pump spins or observer spins. During the actual DEER experiment, the detection pulse sequence generates an electron spin echo signal, referred to as the "refocused echo" (**Figure 6E**). If a dipole-dipole interaction exists between the pump and detection pulse, the pump pulse sequence (blue square in **Figure 6E**) induces oscillations in the refocused echoes, producing the PELDOR/DEER "time trace" (**Figure 6E**). These recorded time traces are then converted into distance distributions (**Figure 6F**) using software such as DeerAnalysis and ComparativeDeerAnalyzer, which apply Tikhonov regularization, or DeerNet, which utilizes a deep neural network trained on simulated DEER datasets to automatically extract distance distributions[86-89].

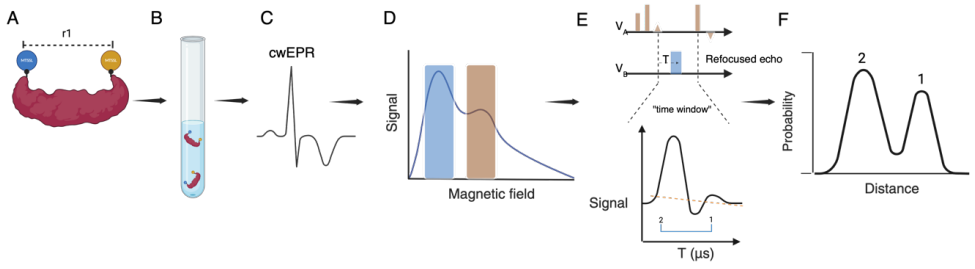


Figure 6. Workflow of measuring of inter-spin distances. (A) Two residues within a detectable distance are selected for spin labeling. (B) Spin labels are attached to two cysteine residues in buffer. (C) cwEPR is used to assess spin labeling efficiency and probe label mobility. (D-F) Workflow of a DEER experiment. Adapted from Lan *et al.* and Peter *et al.*[86,87], and illustrated using BioRender.

1.6.2 X-ray crystallography

X-ray crystallography is a fundamental technique for determining the three-dimensional (3D) structures of proteins at atomic resolution by recording the diffraction of an incident X-ray beam into different directions after hitting the sample[90]. This method provides critical insights into enzyme mechanisms, ligand interactions, and macromolecular conformations, making it valuable for drug discovery, molecular biology, and biophysics. The process begins with protein crystallization, in which protein molecules arrange into a highly ordered periodic lattice. This organization is essential because X-ray diffraction relies on the coherent scattering of photons. Unlike individual protein molecules, which produce weak and incoherent scattering signals, a well-ordered crystal amplifies these scattered signals, significantly improving the signal-to-noise ratio. Additionally, crystallization eliminates the complexities associated with random molecular orientations, which would otherwise complicate data analysis[90].

Successful 3D structure determination *via* X-ray crystallography depends on obtaining well-ordered crystals, where protein molecules adopt a periodic and highly structured arrangement. Various crystallization techniques are employed to achieve this, with vapor diffusion (hanging and sitting drop) and lipidic cubic phase (LCP) methods being among the most widely used[91,92]. Vapor diffusion is commonly applied to all proteins, whereas LCP crystallization is specifically designed for membrane proteins, which require a lipidic environment to maintain their native conformation. The choice of crystallization method depends on factors such as protein solubility, stability, and buffer conditions, often requiring extensive optimization to achieve diffraction-quality crystals[91,92].

Once a well-ordered crystal is obtained, it is subjected to a high-intensity X-ray beam. As the X-rays penetrate the crystal lattice, they interact with the electron clouds surrounding the atoms, causing elastic scattering (**Figure 7A**). While most scattered waves undergo destructive interference and cancel each other out, constructive interference occurs at specific angles, producing a diffraction pattern[93]. This pattern contains critical information about the electron density distribution within the crystal**.** The relationship governing the diffraction pattern is mathematically described by Bragg's Law, which correlates the X-ray wavelength, the angle of the incident beam, and the spacing between lattice planes in the crystal[93].

While x-ray diffraction provides essential structural information, the recorded diffraction patterns alone are incomplete because X-ray detectors capture only the intensities of scattered beams, not their phases—a fundamental challenge known as the phase problem[94,95]. Without phase information, reconstructing an accurate electron density map is impossible, making phase determination a critical step in X-ray crystallography[94,95]. Several approaches have been developed to address this issue. Molecular Replacement (MR) is the most

commonly used method, relying on an existing structural model to estimate initial phases, which are then refined to fit the experimental data[95]. Alternatively, experimental phasing methods, such as Multiple Isomorphous Replacement (MIR) and Single/Multiple-wavelength Anomalous Dispersion (SAD/MAD), involve introducing heavy atoms into the crystal to estimate phase differences[96-98].

Following phase estimation, an electron density map is generated through Fourier transforms, representing the spatial distribution of electrons in the crystal (**Figure 7B**). Model building involves fitting atomic coordinates into the electron density map, a process facilitated by COOT[99]. Refinement algorithms are then applied to minimize discrepancies between the model and the experimental data. Programs such as PHENIX optimize atomic coordinates, B-factors (atomic displacement parameters), and occupancy values, improving overall structural accuracy (**Figure 7C**)[100]. The final structure is validated using metrics such as the Ramachandran plot, which ensure stereochemical accuracy and structural quality[101].

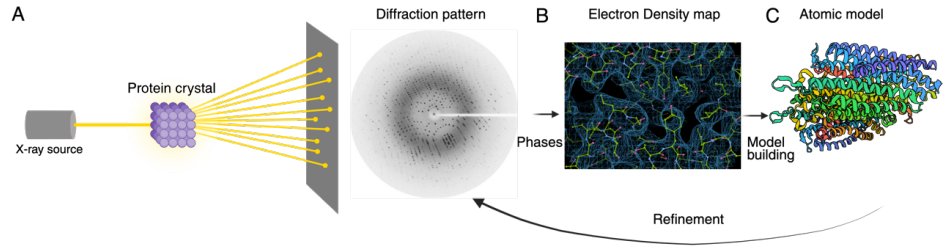


Figure 7. Workflow of solving a protein structure using crystallography. (A) Crystallized protein is exposed to X-ray radiation, and the diffraction pattern is captured, containing intensity data. (B) Intensity-based amplitude data is combined with phase information to generate an electron density map. (C) A structural model is built into the electron density, creating an atomic-level representation of the macromolecule. Throughout the process, the model is iteratively refined to the original diffraction data to minimize overfitting. Adapted from Srinivasan[102], illustrated using BioRender.

1.6.3 Cryogenic electron microscopy

Cryo-electron microscopy (Cryo-EM) has become a transformative technique in structural biology, enabling researchers to determine the three-dimensional structures of biological macromolecules at atomic resolution. Its ability to circumvent the need for crystallization and visualize samples in their near-native states has made it an valuable tool for studying dynamic protein complexes, membrane proteins, and large molecular assemblies[103]. Cryo-EM operates on the principle of electron-matter interactions. A high-energy electron beam, typically operating at 200–300 kV, is transmitted through a biological sample, and the scattered electrons are recorded to generate a series of 2D projections[104]. Unlike X-ray crystallography, which requires crystalline lattices, Cryo-EM visualizes individual molecules[104]. This technique has revolutionized structural biology, particularly for large macromolecular complexes and membrane proteins that are difficult to crystallize. However,

biological samples are highly sensitive to radiation damage, necessitating the use of low-dose electron beams. To reduce thermal motion and dehydration, the sample is embedded in vitrified ice—a non-crystalline form of water—by rapid freezing in liquid ethane at cryogenic temperatures (~77 K)[105,106]. This vitrification process prevents the formation of ice crystals, preserving the structure of the molecules in a native-like environment[105].

Optimizing Cryo-EM sample preparation is critical for achieving high-resolution reconstructions and depends on factors such as protein quality, buffer composition, and grid selection (**Figure 8**)[107]. Significant advances have been made to ensure uniform sample distribution on grids, particularly for challenging targets like membrane proteins[108]. Strategies include using continuous carbon or graphene-oxide grids, as well as adding secondary detergents to mitigate the effects of the air-water interface[108]. However, there is no universal protocol suitable for every protein. Each sample requires tailored optimization to address its unique properties and ensure the highest-quality data is obtained.

During Cryo-EM data collection, scattered electron waves produce images with inherently low contrast and signal-to-noise ratio (SNR)[109,110]. To overcome these limitations, thousands of particle images must be computationally averaged to reconstruct the 3D structure using algorithms based on Fourier transforms and single-particle analysis[111]. Following Cryo-EM data collection, multiple computational steps are performed to enhance the signal and reconstruct the structure (**Figure 8**)[112]. The workflow begins with motion correction, where algorithms such as MotionCor2 compensate for beam-induced motion, aligning individual frames from dose-fractionated movies to stabilize the images[113]. This is followed by contrast transfer function (CTF) estimation, which corrects for phase distortions caused by the lens aberrations and determines defocus values, quantifying the focal plane displacement of electron beams relative to the specimen plane[114]. Given that biological samples have weak electron scattering contrast, defocus enhances image visibility[114]. Software tools such as CTFFIND4, RELION and CryoSPARC are commonly used for CTF estimation and correction, improving the accuracy of downstream structural analysis[115-117]. Particle picking is then performed to extract individual particle images from the micrographs. These particle images undergo 2D classification, where they are grouped based on similar orientations, generating averaged 2D class images that improve the SNR. The aligned particle images are subsequently used for 3D reconstruction by employing computational algorithms such as RELION or CryoSPARC[115,116].

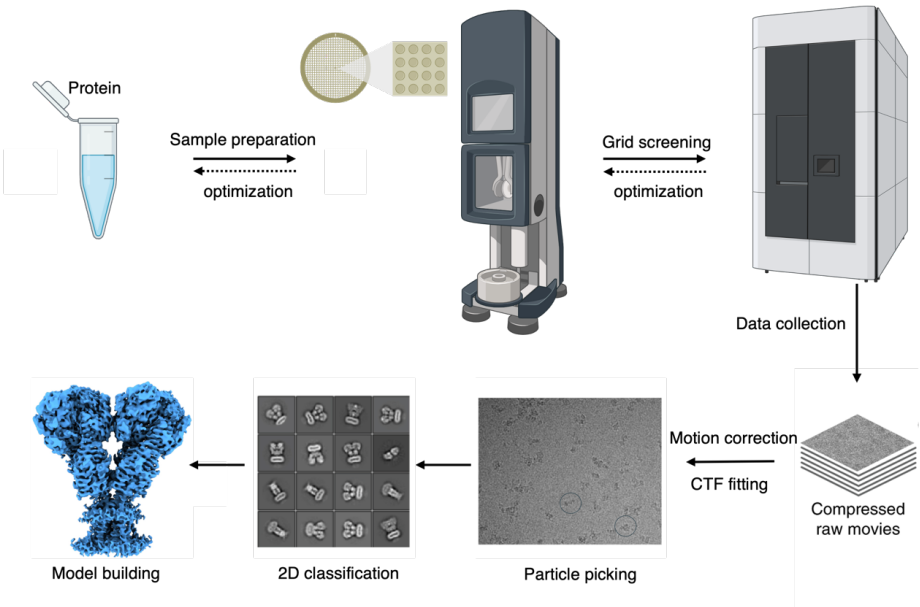


Figure 8. Workflow of the cryo-EM sample preparation and structure determination. Illustrated using BioRender

1.6.4 Nanion SURFE2R N1

The Nanion SURFE²R N1 is an advanced electrophysiology platform designed to measure the activity of electrogenic transporters, ion pumps, and other membrane proteins with exceptional sensitivity. Developed by Nanion Technologies, it utilizes solid-supported membrane (SSM)-based electrophysiology[118,119] (**Figure 9A**), a highly specialized technique that enables the precise detection of charge movements associated with transporter function. This approach provides a label-free, high-sensitivity alternative to traditional electrophysiological methods, making it ideal for studying ion channels, transporters, pores and toxins[120-122].

unlike patch-clamp, which measures stationary currents by clamping voltage, SSM-based electrophysiology measures transient currents driven by a substrate gradient[123]. As charged molecules move, they generate a membrane potential, which is detected *via* capacitive coupling (**Figure 9B**). Transport ceases when these potential reaches equilibrium with the chemical gradient, causing the current to become transient[118,119].

Proper sample preparation is essential for achieving reliable results with the Nanion SURFE²R N1. Purified proteins are incorporated into synthetic lipid bilayers and then removing the detergent *via* dialysis or bio-beads [124], allowing for the self-assembly of proteoliposomes that mimic the native membrane environment. Once the sample is

prepared, it is applied to the gold-coated sensor chip for immobilization (**Figure 9C**). The sensor surface is activated to allow stable adsorption of membrane vesicles or proteoliposomes, ensuring uniform attachment for optimal current detection[118].

The data collection process begins with an initial baseline measurement, in which the system records the resting signal of the sensors before applying substrates. Upon applying the substrate, the transporter undergoes conformational changes, moving charged molecules across the membrane and generating a transient current (**Figure 9D**). The recorded signals are automatically digitized and stored, ensuring precise and reproducible measurements.

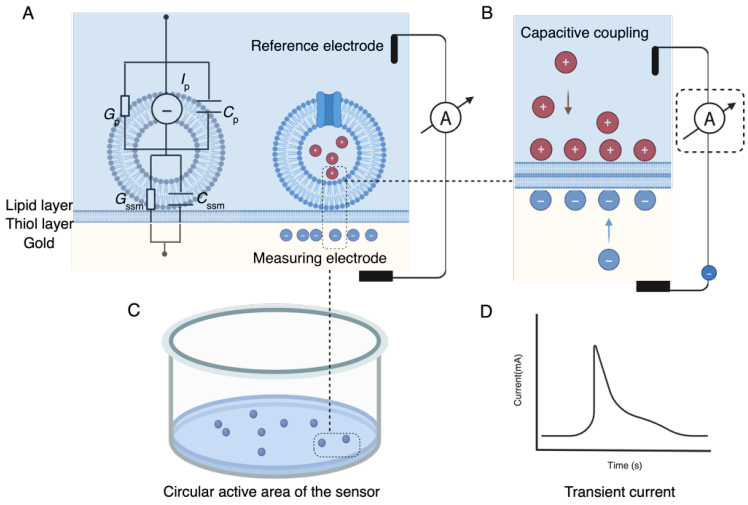


Figure 9. The principles of SSM-based electrophysiology. (A) Proteoliposomes are adsorbed onto the solid-supported membrane (SSM), forming a capacitively coupled system. Both the proteoliposomal membrane (P) and the SSM can be characterized by their respective capacitance (C) and conductance (G). Electrogenic transport leads to the accumulation of charge within the proteoliposomal membrane, which in turn induces a corresponding charge displacement at the measuring electrode, enabling detection of transport activity. (B) A close view of the interface of the SSM and the proteoliposomes. The accumulation of charge within the proteoliposomes generates a measurable current. (C) Proteoliposomes bound to the SSM contribute to the overall current signal. (D) Representative transient current responses observed in the output data. Adapted from Bazzone *et al.,*[125], and illustrated using BioRender.

2 Aim of Study

The focus of this project was to gain a deeper understanding of the mechanism in mPPases and their conformational changes during the hydrolysis process. Additionally, the project aimed to identify inhibitors targeting PfPPase using a combination of biochemical and biophysical approaches.

The more specific aims were:

1. To understand the catalytic asymmetry and the "pumping-after-hydrolysis" mechanism of mPPase in crystal and physiological conditions (study I and II).
2. To establish and characterize the expression and purification of PfPPase and its enzymatic activity (study III).
3. To identify the nonphosphorus inhibitors with the potential for specific inhibition of TmPPase and PfPPase (Study III and IV).

3 Materials and Methods

DNA constructs (Table 5), cell lines (Table 6), experimental methods (Table 7) and Software or programs (Table 8) that used in this thesis are listed in the table below. Full details are given in papers I, II, III and IV.

Table 6. List of constructs used in the thesis project.

Construct	Vector	Study
TmPPase (WT)-His_6	pRS1024	I-IV
TmPPase (T211C)-His_6	pRS1024	I
TmPPase(S525C)-His_6	pRS1024	I
PfPPase-GS_{10}-TEV-eGFP-His_8	pK503.9	III
PfPPase-GS_{10}-TEV-His_8	pK503.9	III, IV

Table 7. List of cell lines used in the thesis project

Cell line	Manufacturer	Culture media	Study
Saccharomyces cerevisiae strain BJ1991		Yeast extract peptone dextrose (YPD) medium	I-IV
Spodoptera frugiperda (Sf9)	Thermo Fisher Scientific	XPRESS medium (Lonza)	III
Trichoplusia ni High Five (Hi5)	Thermo Fisher Scientific	XPRESS medium (Lonza)	III, IV

Table 8. List of experimental methods used in the thesis project

Experimental method	Study
Protein expression and purification (TmPPase)	I-IV
Protein expression and purification (PfPPase)	III, IV
Membrane extraction and detergent screening	III
TmPPase activity assay	I, II
PfPPase activity assay	III, IV
In-gel activity assay	III
Crystallisation and structure determination	I, II

EPR spectroscopy	I
Double Electron-Electron Resonance (DEER, or PELDOR) spectroscopy	I
Electrometric measurement	I, II
SDS PAGE	I-IV
Native PAGE	III
SEC-MALS	III
Inhibitor screening	I, III, IV

Table 9. List of programs used in the thesis project

Program	Reference	Study
Phaser: Molecular replacement	[126]	I
Phenix: Model refinement	[100]	I
Coot: Model building	[99]	I
Pymol: Visualization of structures	The PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.	I
DeerAnalysis: DEER data analysis	[127,128]	I
ComparativeDeerAnalyzer: DEER data analysis	[128]	I
MttslWizard: Prediction of distance distribution between two sites	[83]	I
GraphPad Prism 10: Data analysis and curve fitting	Prism 10	I-IV
R3.6.3 and n-parameter logistic regression (nplr) package: Determination of IC_{50} values	[129], [130]	III, IV
ImageJ: PAGE analysis	[131]	III

3.1 Methods in structural studies of PfPPase

3.1.1 Protein expression and purification

I designed two constructs for the PfPPase-BRIL fusion protein, one with a 32-amino-acid linker between PfPPase and the C-terminal BRIL (PfPPase-BRIL_C), and the other with an 18-amino-acid linker between PfPPase and the N-terminal BRIL (PfPPase-BRIL_N). Both constructs include a tobacco etch virus (TEV) protease cleavage site and a His_{10}-tag. The cloning, transformation, expression, and purification procedures were performed as described in **Study III.** Briefly, His-tagged PfPPase_BRIL in pK509.3 plasmids were transformed into *E. coli* DH10EMBacY and subsequently transfected into Hi5 insect cells. Expression was harvested 48 h post-infection. PfPPase-BRIL was solubilized in DDM+CHS and purified in GDN using immobilized metal affinity chromatography (IMAC).

The expression and purification of BAK5 followed the protocol developed by Mukherjee *et al*.,[132]. Plasmids encoding the BAK5 synthetic antibody were kindly provided by Maria Nikolova *via* material transfer agreement with Mukherjee *et al.,*[132]. *E. coli* BL21 Star™ (DE3) were transformed with the plasmids, and cells were cultured in Luria-Bertani (LB) medium supplemented with 100 μg/ml carbenicillin. Protein expression was induced with 0.4 mM IPTG when the culture reached an OD_{600} of 0.8. Cells were harvested by centrifugation at 6000 × g for 30 mins at 7 °C.

Harvested cells were lysed using an EmulsiFlex C3 (Avestin) in 30 mL of lysis buffer containing 20 mM HEPES-NaOH (pH 7.4), 500 mM NaCl, 1.33 mM DTT, 1.67 mM PMSF, and one tablet of Pierce protease inhibitors (Thermo Scientific™). Cell debris was removed by centrifugation at 6000 × g.

BAK5 protein was purified using a Protein L column. The column was equilibrated with 20 column volumes (CV) of water, followed by 10 CV of running buffer (20 mM HEPES, pH 7.4, and 500 mM NaCl). The lysate was loaded onto the column, washed with 20 CV of running buffer, and eluted with 0.1 M acetic acid and collected in 2 mL fractions. Each fraction was transferred into tubes pre-loaded with 40 μL of 1 M HEPES-NaOH (pH 8.0), adjusting the final buffer concentration to 20 mM HEPES-NaOH (pH 8.0). BAK5 was then concentrated to a maximum of 5 mg/mL using Amicon centrifugal concentrators (10 kDa molecular weight cut-off, Millipore EMD), frozen in liquid nitrogen, and stored at -80 °C.

3.1.2 Formation of the PfPPase-BRIL/BAK5 complex

BAK5 was thawed and mixed with freshly purified PfPPase-BRIL at a 1.5:1 molar ratio, followed by incubation on ice for 2 h. The mixture was centrifuged at 25,000 × g for 30 mins at 4 °C to remove any precipitates. The SEC buffer contained 20 mM MES (pH 6.5), 50 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, and 0.006% GDN. The SEC column was pre-equilibrated with the SEC buffer before sample loading. The sample was separated using either a Superdex 200 Increase 5/150 GL or a Superdex 200 Increase 10/300 GL column. The flow rate was set at 0.05 mL/min for the 5/150 column and 0.25 mL/min for the 10/300 column Eluted protein factions were collected in 40 μL aliquots for the 5/150 column and 400 μL aliquots for the 10/300 column, followed by SDS-PAGE analysis.

3.1.3 Negative stain electron microscopy (EM)

Negative stain grids (300-mesh copper grids coated with carbon) were glow-discharged using a Pelco Easiglow discharge unit at 10 mA for 30 s. A 5 μl Droplets of the protein sample (0.01 mg/mL) was applied to each grid and incubated for 1 min. Excess sample was quickly blotted off with filter paper, and the grids were washed twice with 10 μL drops of

deionized water, followed by additional blotting. The grids were then stained with 8 µL of 1% uranyl acetate solution for 30 s before being blotted and air-drying for 3 mins at RT.

Grids were screened using a TEM Hitachi HT7800 microscope at 200 kV. Micrographs were manually collected at 50, 000x magnification. CTF correction and particle picking were performed manually in Scipion 3, followed by 2D classification using the same software[133].

3.1.4 Cryo-EM sample preparation

Cryo-EM grid selection was optimized using various grid types, including Quantifoil R 1.2/1.3 300 gold mesh, Au-flat R 1.2/1.3 300 gold mesh, C-flat R 1.2/1.3 gold mesh, graphene oxide on Quantifoil R 1.2/1.3 copper mesh, and Quantifoil R 1.2/1.3 Plus C2 on 300 copper mesh grids. Glow discharge was performed using either a Pelco Easiglow discharge unit (10 mA, 30 s) or *via* in-line glow discharge within the VitroJet system**.** For grid preparation, 50 nL of a 0.5 mg/mL protein sample was applied to the glow-discharged grids, and vitrified using the VirtroJet system, operated by Behnam Lak (University of Helsinki, Finland). Additionally, 3 µL of protein samples at 2.5 mg/mL was mixed with 3 mM fluorinated Fos-Choline-8 (Anatrace) immediately before apllication to Au-flat R 1.2/1.3 300 gold mesh grids. These grids were prepared using the Vitrobot Mark IV (Thermo Fisher) with a 6 s blotting time at 4°C with 80 % humidity.

3.1.5 Data acquisition and initial cryo-EM data processing

For grid screening, small-scale data collection was performed on a FEI TALOS Artica electron microscope equipped with a Falcon 4i direct electron detector camera. Data were acquired at a nominal magnification of 150,000×**,** corresponding to a pixel size of 0.97 Å/pixel**.** A total of 100 micrographs per grid were collected within a defocus range of –1 to –3 µm**,** with a cumulative electron dose of 40 e^{-}/Å². Based on these results, the optimal Au-flat R 1.2/1.3 300 gold mesh grid was selected for large-scale data collection. High-resolution data acquisition was conducted using a Titan Krios electron microscope equipped with a Falcon 4i direct electron detector operating in integrative mode at a resolution of 0.74 Å/pixel and a nominal magnification of 165,000×. A total of 12,000 movies were acquired with a defocus range of –1.2 to –2.4 µm, a per- fraction dose of 1 e^{-}/Å², and a total cumulative dose of 40 e^{-}/Å².

Cryo-em data processing was carried out using CryoSPARC 4.6, an integrated software platform for single-particle cryo-em analysis[116]. Motion correction, CTF estimation, particle picking, and 2D classification were all performed using CryoSPARC 4.6, ensuring an efficient and streamlined workflow.

4 Results and Discussion

4.1 Catalytic asymmetry in mPPase

We previously reported the asymmetric time-resolved crystal structures of TmPPase in the presence of PP_i[48]. In this study, we aim to determine whether the conformational changes occurring during hydrolysis in a native-like environment are consistent with the asymmetric mode observed in crystal structures. To address this question, we employed PELDOR spectroscopy to measure the inter-spin distances in TmPPase in solution. Capturing these conformational changes in the presence of the PP_i presents a significant challenge, as the cryo-trapping approach used in time-resolved crystallography is incompatible with PELDOR measurements. To circumvent this limitation, we optimized the use of PP_i analogs, such as bisphosphonates, to mimic key stages of enzymatic hydrolysis and stabilize potential asymmetric conformations for analysis.

4.1.1 TmPPase structures in complex with bisphosphonate

We obtained the crystal structures of TmPPase in complex with etidronate (ETD) and zoledronate (ZLD) at resolutions of 3.2 Å and 3.3 Å, respectively. These structures provide insights into the conformational changes induced by bisphosphonate binding.

The overall structure of the TmPPase:ETD closely resembles the resting state TmPPase:Ca structure (PDB ID: 4AV3) with a RMSD/Cα of 0.46 Å (**Study I, Table S2**). However, unlike the fully open conformation, the TmPPase:ETD structure exhibits extra electron density above the hydrolytic center in both monomers, allowing for partial modeling of loop5-6 residues (**Figure 10A**). Structural superposition of monomers A and B highlights significant asymmetry, with an RMSD of 1.44 Å (**Study I, Table S2**), which is higher than the 0.39 Å RMSD observed between monomers in the resting state (**Study I, Table S2**). Monomer B adopts a more constricted conformation, characterized by inward bending of inner ring helices (TMH 11, 12, and 15) towards the active site and deeper insertion of cytoplasmic loops (loops11-12, 13-14, and 15-16) into the active site (**Figure 10A**).

At the active site, ETD binding induces and stabilizes the movement of loop5-6 (**Figure 10B**), a feature also observed in the TmPPase:600s structure (**Figure 4E**). Notably, the distinct orientations of ETD in the two active sites contribute to conformational differences

between monomers A and B (**Figure 10B**). Two phosphonate groups separated by a central hydroxyl-substituted carbon. In both monomers, the lower phosphonate group aligns with the leaving-group phosphonate of IDP. However, the upper phosphonate groups adopt different orientations in monomer A. ETD_A is tilted by ~35.9° relative to IDP, whereas in monomer B, ETD_B remains nearly parallel. Consequently, ETD_A and ETD_B exhibit distinct interaction profiles within the loop5-6 and the active site. In monomer A, loop5-6 extends toward the active center, where it interacts with ETD_A *via* $E^{5.76}$, mediated by a Mg^{2+} ion (**Figure 10C**). In contrast, in monomer B, loop5-6 interacts with ETD_B *via* $D^{5.77}$, also mediated by Mg^{2+} (**Figure 10D**). Additionally, in monomer B, $D^{11.57}$, $D^{12.39}$, and $N^{12.43}$ within TMH 11 and TMH 12 interact with ETD_B *via* a water molecule, inducing a subtle inward shift of these helices. This structural rearrangement results in a more constricted conformation in monomer B.

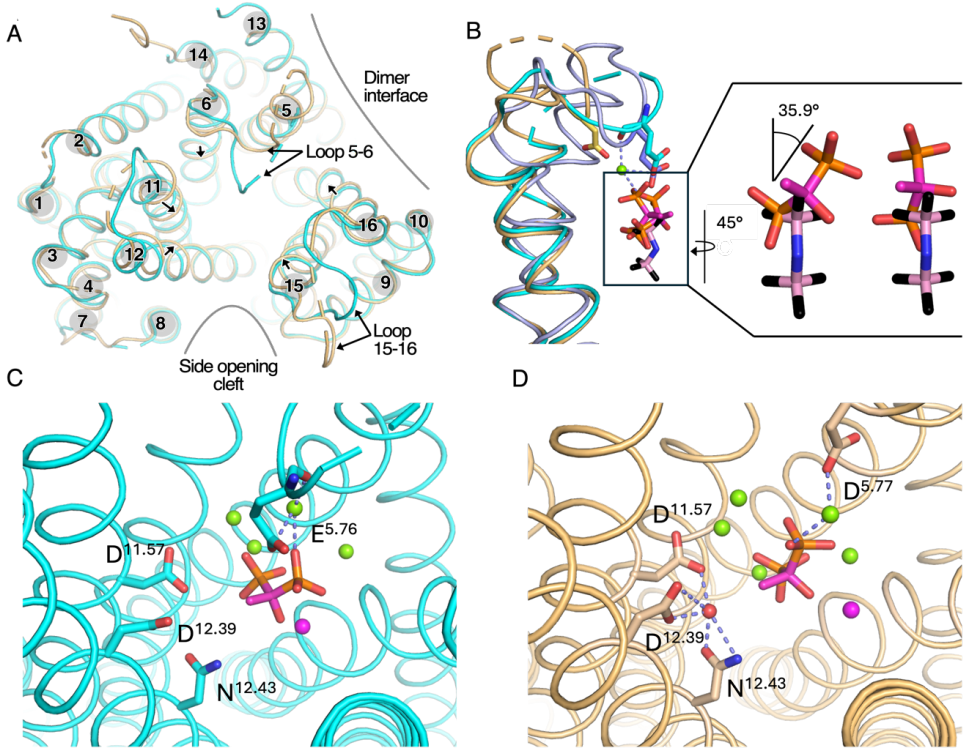


Figure 10 Structural asymmetry in TmPPase:ETD complex. (A) Superposition of chain A (cyan) and chain B (wheat) showing the relative movements (black arrow) of helices. (B) Superposition of TMH5 and TMH6 from TmPPase:ETDs and TmPPase:IDP (light blue; PDB: 5LZQ) showing the movement of loop5-6 and reorientation of ETD_A, ETD_B and IDP. Dashed lines show the interaction of $E^{5.76}$ of loop5-6 in Chain A with ETD_A.; Close-up view of IDP superposed with ETD_A and ETD_B. (C) Residues in the active site with ETD_A coordinated (dashed lines), Ca^{2+} (pink sphere) and nucleophilic water (red spheres) in a Mg^{2+} metal cage (green spheres). (D) Residues in active site with ETD_B coordinated (dashed lines), Ca^{2+} (pink spheres) and water (red spheres) in a Mg^{2+} metal cage (green spheres).

In contrast, the TmPPase:ZLD structure adopts a more closed conformation in both monomers, with an RMSD/Cα of 0.51 Å (**Study I, Table S2**), closely resembling the TmPPase:IDP structure (RMSD/Cα of 0.76 Å). The most notable distinction between these structures lies in the active site, where the TmPPase:ZLD structure does not achieve fully closure (**Figure 11A**). Specifically, the tunnel extending from the hydrolytic centre to the cytoplasmic opening remains open, in contrast to the fully sealed tunnel observed in the IDP-bound structure (**Figure 11B**). This discrepancy is likely arttributed to the bulkier heteroaryl group of ZLD_S, which induces an outward displacement of TMH 11, 12, and 15 on the cytoplasmic side (**Figure 11C-D**), thereby preventing complete closure at the active site. Moreover, although the arrangement of TMHs in the ion gate is structurally similar to that in TmPPase:IDP (**Study I, Fig. EV7C**), no electron density corresponding to Na^+ was detected in the ion gate of the TmPPase:ZLD structure. This absence is particularly notable given that the TmPPase:ZLD structure was resolved at a higher resolution (3.26 Å) compared to the IDP-bound structure (3.5 Å).

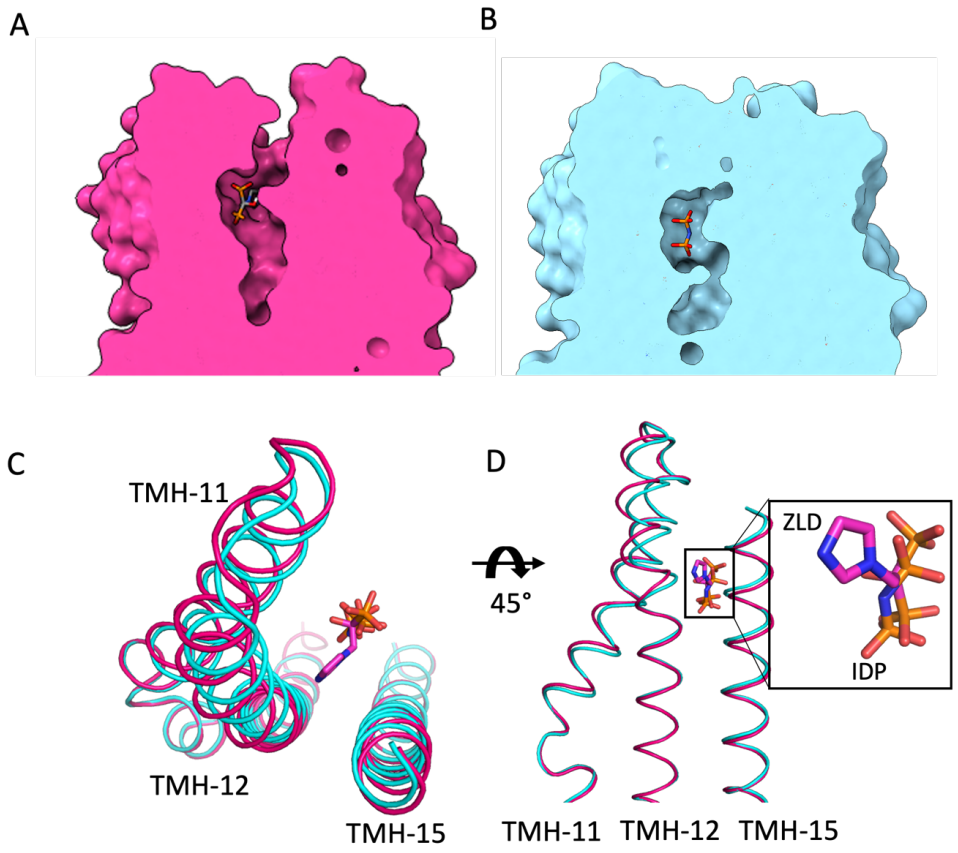


Figure 11. Comparison of the TmPPase:ZLD and TmPPase:IDP structures in the active site. (A) Cross-section view of the active site in TmPPase:ZLD (pink). (B) Cross-section view of the active site in TmPPase:IDP (cyan). (C) Superposition of TMH11, TMH12 and TMH15 in TmPPase:ZLD and TmPPase:IDP showed the movement of the hydrolytic centre and the orientation of ZLD and IDP.

Overall, the two newly solved bisphosphonate-bound TmPPase structures adopt a conformation that is neither fully open nor closed, differing from those obtained under other conditions, such as with IDP or Ca^{2+}. The TmPPase:ETD structure further supports the asymmetric binding mode in TmPPase. However, ETD is positioned above the hydrolytic center (**Figure 10B**) and is unable to descend further due to the presence of Ca^{2+} in the active site (**Figure 10C, D**), similar to what is observed in the TmPPase:Ca structure (PDB ID: 4AV3). As a result, although ETD induces an asymmetric conformation, partial closure of loop5-6, and contributes to some stabilization, the overall arrangement of the inner and outer helices remains more closely resembles the open state rather than the fully closed conformation. To investigate how TmPPases behave under more physiologically relevant conditions, we utilized DEER spectroscopy to measure conformational dynamics in solution, both in the absence/presence of bisphosphonates.

4.1.2 Determination of dynamic changes in TmPPase in solution

To track conformational changes across the membrane without compromising enzymatic activity, we introduced cysteine mutations at three selected sites: S525 (located at the periplasmic side), C599, and T211 (both located at the cytoplasmic side). Dr. Anokhi Shah and coworkers spin-labeled them with MTSSL, denoted as R1 (C599R1, S525R1, and T211R1), to monitor diploe-dipole interactions between the MTSSL labels. In addition, we generated *in silico* distance distribution predictions for C599R1, S525R1, and T211R1 using MtsslWizard[83], based on TmPPase crystal structures bound to different ligands (**Figure 12**). Here, we refer to the conformation of the TmPPase:IDP structure as 'closed' at both sides, even for residues not in the active site, for residues as in the TmPPase:Ca structure as 'open' at both sides. Our asymmetric model has, for instance, S525(A) 'closed' but S525(B) 'open'.

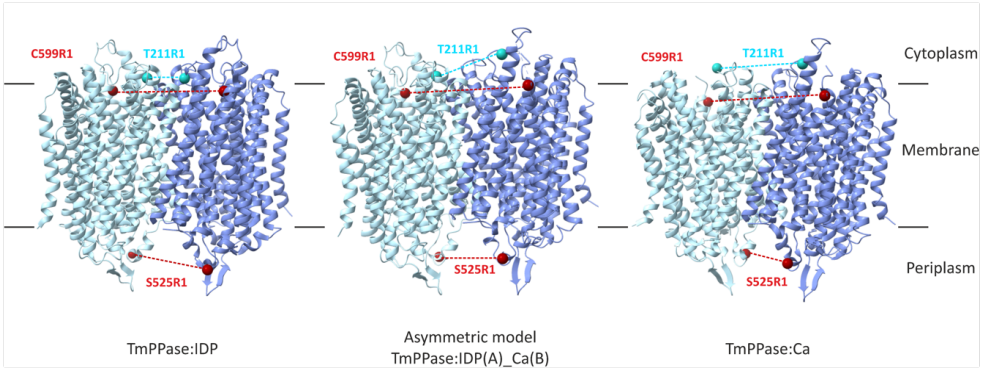


Figure 12. Predicted distance distribution for labeled sites. Symmetric structures (TmPPase:IDP (PDB: 5LZQ), asymmetric model (TmPPase:IDP(A)_Ca(B)), and TmPPase:Ca (PDB: 4AV3) of TmPPase. The sites mutated and labeled with MTSSL are shown as spheres: T211R1 (cyan) and C599R1 (maroon) are located on the cytoplasmic side (top), and S525R1 (maroon) is on the periplasmic side (bottom) of the membrane. Distances between spin-labeled pairs are indicated by dashed lines, matching the color of the corresponding spheres.

C599R1 dimer generated high-quality DEER trances under all eight conditions (*apo*, +Ca, +Ca/ETD, +ETD, +IDP, +ZLD, +Pamidronate (PAM), +Alendronate (ALE))(**Figure 13A**). The prominent dipolar oscillations observed in the raw DEER data were background-corrected using DeerAnalysis2022[88] (**Figure 13B**) and then converted into distance distribution. These distributions were overlaid with predicted ones derived from solved structures (**Figure 13C**). Across all conditions, the modal DEER distance distribution is approximately 5.8 nm, which is longer than the predicted 4.8 nm distance derived from the TmPPase:IDP structure – where both monomers are closed, and significantly shorter than the predicted 6.6 nm distance for the TmPPase:Ca and TmPPase:ETD structures, in which both monomers are open. Instead, the measured 5.8 nm distance aligns closely with the asymmetric model, termed TmPPase:IDP(A)_Ca(B), supporting an asymmetric closed-open state of TmPPase on the cytoplasmic side in solution. A shoulder and small peak at ~5 nm, seen in the presence of IDP and ETD, are consistent with the predicted distance for the TmPPase:IDP structure (**Figure 13C**), where both monomers are in a fully closed conformation. We interpret this as representing a population in which IDP or ETD binds to both active sites, leading to a closed–closed conformation at the cytoplasmic side. This interpretation is consistent with structural data and the mechanism of substrate inhibition[48,59]. In this scenario, the substrate occupies the second active site, reducing enzymatic activity and promoting a fully closed state in both monomers. Of course, we cannot completely exclude the possibility that the monomers adopt an intermediate state, leading to a distance distribution similar to that of the hybrid asymmetric structure under *apo*, +Ca, and +Ca/ETD conditions, but it is not consistent with the presence of shoulders in the IDP and ETD data, and with the major peak seen in the IDP, ETD and ZLD data.

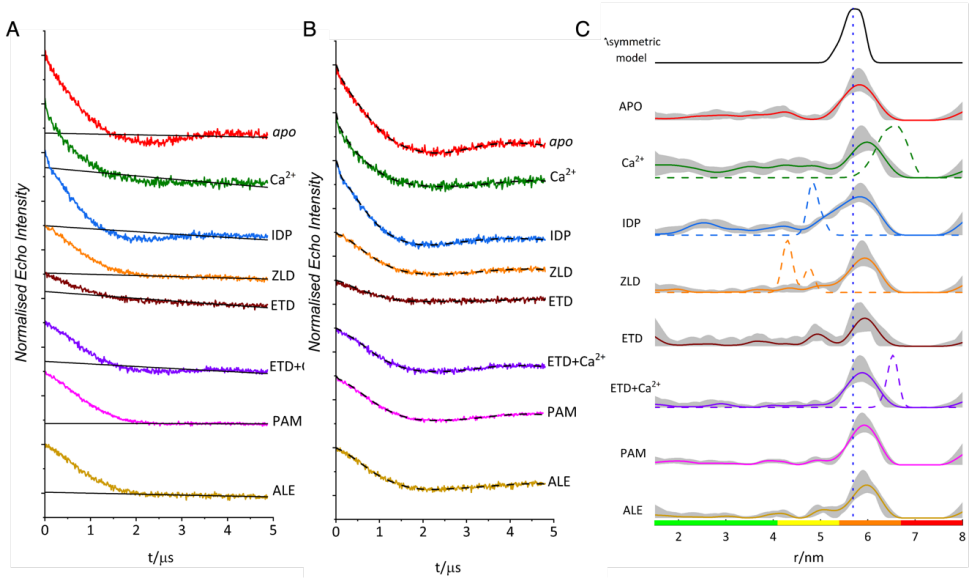


Figure 13. DEER distance distributions of TmPPase S525R1 under different conditions. (A) DEER raw data traces for C599R1. Each condition is labelled, and the raw data are colour-coded, with the background

function indicated as solid black lines. (B) DEER background-corrected time-domain traces for C599R1, respectively. The vertical black dashed line represents the minimum of the first oscillation in the *apo* state and aids visualisation to highlight the shifts in the oscillation minimum under different conditions. (C) Distance distributions of C599R1. The *in-silico* distance distribution corresponding to each spin pair modelled onto the asymmetric hybrid structure (TmPPase:IDP(A)_Ca(B)) is shown at the top as a solid black line, with the modal distance shown as a vertical dashed line. *In silico* predicted distance distributions for each condition, modelled using the solved structures (TmPPase:Ca, TmPPase:ETD (PDB 9G8K), TmPPase:ZLD (PDB 9G8J), and TmPPase:IDP) are presented as coloured dashed lines overlaying the experimental distributions. The shaded regions represent the 95% (2σ) confidence interval of the distributions, and the colour bars represent an assessment of the reliability of the distributions. The probability density within the green region indicates the mean distance, width, and peak shape are all reliable; the probability density within the yellow region indicates the mean distance and width are reliable; the probability density within the orange region indicates that the mean distance is reliable; the probability density within the red region indicates no quantitation is possible.

The S525R1 dimer also produced high-quality DEER traces, enabling precise resolution of inter-spin distance variations under different inhibitor conditions (**Figure 14A**). Although the observed changes were small, they were statistically significant in the raw data[82,85,87,134]. Background correction of the DEER traces was performed using DeerAnalysis2022[88] (**Figure 14B**). In the *apo* state, the modal distance was measured at 3.8 nm with broad distribution, which decreased to 3.6 nm upon the addition of Ca^{2+}, consistent with the predicted distances derived from the TmPPase:Ca structure (**Figure 14C**). When both Ca^{2+} and ETD were present, the modal distance increased slightly to 3.7 nm, aligning with predictions for TmPPase:ETD structure. The addition of ETD alone (without Ca^{2+}) resulted in an increased modal distance of 3.9 nm (no structure is available for ETD only condition). IDP induced the increasing of the modal distance distribution to 4.0 nm, shorter than the predicted value based on the TmPPase:IDP structure (4.3 nm). In contrast, ZLD produced the shortest modal distance (3.4 nm), significantly deviating from the prediction generated by MtsslWizard (4.3 nm). Meanwhile, with the addition of PAM and ALE, the resulting distributions show modal distances of 4.1 nm and 4.3 nm (no structure available), respectively. Since we do not have structures for their complexes with TmPPase, their orientation in the active site remains unknown.

Given the narrower distance range from 3.6-4.0 nm, which resembles both asymmetric (*i.e.* closed-open) and *apo*/Ca^{2+} state (*i.e.* open-open) models at the periplasmic side. We employed Bhattacharyya coefficients[135] to evaluate which model best fits the S525R1 DEER data. The values are as follows: +Ca = 0.98 (*apo*), 0.90 (asymmetric model); +IDP = 0.97 (*apo*), 0.98 (asymmetric model); +ETD = 1.0 (*apo*), 0.97 (asymmetric model); +ZLD = 0.95 (*apo*), 0.84 (asymmetric model); +Ca/ETD = 0.98 (*apo*), 0.91 (asymmetric model). The calculation of Bhattacharyya coefficients for the 525R1of +PAM and +ALE conditions are not available, as the data were recorded on a different instrument with a non-comparable x-axis. These coefficients for S525R1 indicate that the *apo*-state (*i.e.* open-open) model provides a better fit to the experimentally derived distributions for +Ca, +Ca/ETD, +ETD, and +ZLD. In contrast, the asymmetric (*i.e.* closed-open) model better describes the experimental data for +IDP. Higher Bhattacharyya coefficient values (closer to unity) indicate greater overlap, which here is used as a proxy for model agreement.

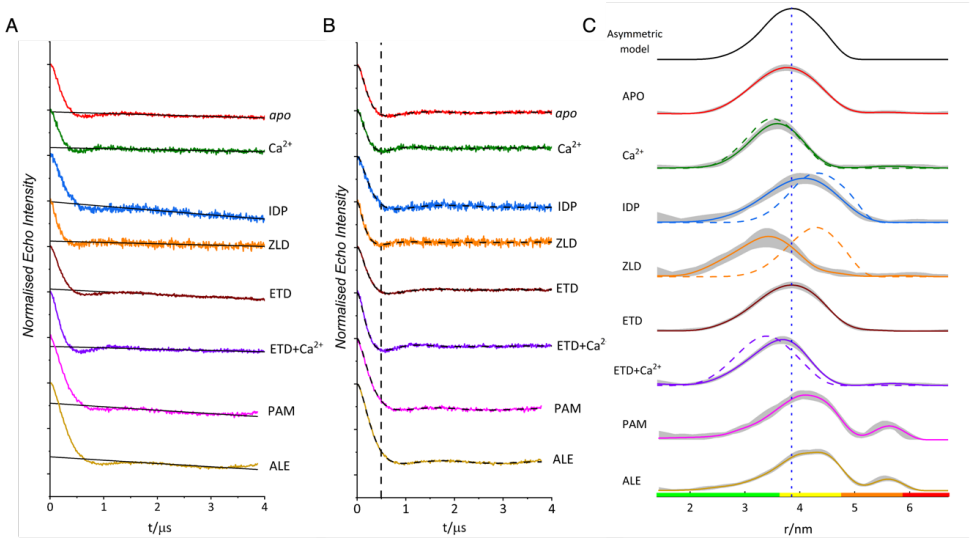


Figure 14. DEER distance distributions of TmPPase S525R1 under different conditions. (A) DEER raw data traces for S525R1. Each condition is labelled, and the raw data are colour-coded, with the background function indicated as solid black lines. (B) DEER background-corrected time-domain traces for S525R1, respectively. The vertical black dashed line represents the minimum of the first oscillation in the *apo* state and aids visualisation to highlight the shifts in the oscillation minimum under different conditions. (C) Distance distributions of S525R1. The *in-silico* distance distribution corresponding to each spin pair modelled onto the asymmetric hybrid structure (TmPPase:IDP(A)_Ca(B)) is shown at the top as a solid black line, with the modal distance shown as a vertical dashed line. *In silico* predicted distance distributions for each condition, modelled using the solved structures (TmPPase:Ca, TmPPase:ETD (PDB 9G8K), TmPPase:ZLD (PDB 9G8J), and TmPPase:IDP) are presented as coloured dashed lines overlaying the experimental distributions. The shaded regions represent the 95% (2σ) confidence interval of the distributions, and the colour bars represent an assessment of the reliability of the distributions. The probability density within the green region indicates the mean distance, width, and peak shape are all reliable; the probability density within the yellow region indicates the mean distance and width are reliable; the probability density within the orange region indicates that the mean distance is reliable; the probability density within the red region indicates no quantitation is possible.

Unlike 525R1, the lack of oscillations in the raw DEER data (**Figure 15A**) and increased spin label mobility observed by CW-EPR (**Study I, Figure EV9A**) for T211R1 resulted in broad distributions under all conditions (**Figure 15B-C**), reflecting the high mobility of loop5-6. This observation is consistent with the absence of partial density for loop5-6 in structural data (**Figure 10A**). The *apo*, Ca^{2+}, and Ca^{2+}/ETD conditions showed similar distributions, while the addition of ETD or IDP resulted in shorter modal distances, indicating the closure of the active site. However, these distances were longer than the predicted distance in TmPPase:IDP structure, where both active sites are fully closed. In the presence of ZLD, the distribution shows a contribution from longer distances within the conformational ensembles, deviating from the predicted distances in the TmPPase:ZLD structure. Due to the featureless raw DEER data recorded for T211R1 and the resulting broad distance distributions, we refrain from further interpreting conformational changes based on this mutant.

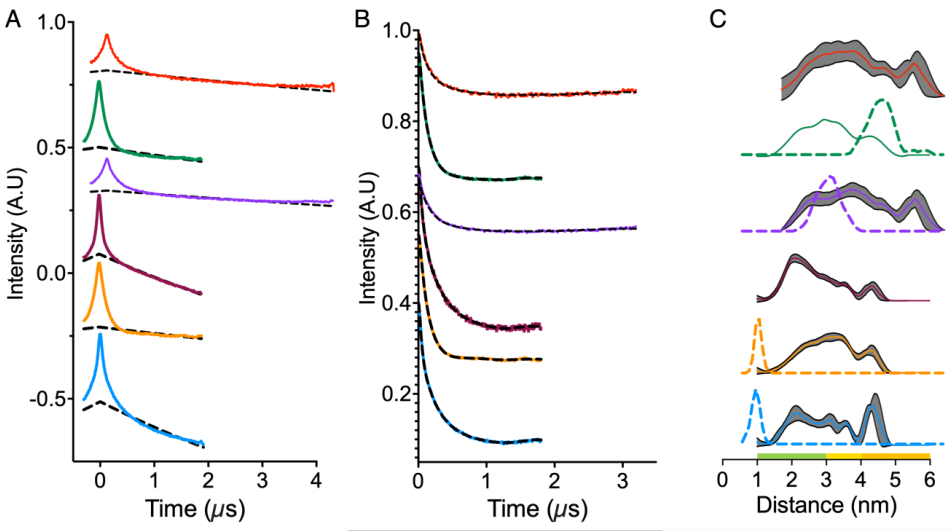


Figure 15. DEER distance distributions of TmPPase T211R1 under different conditions. (A) DEER raw data traces for T211R1. Each condition measured is coloured according to the condition used. (B). DEER background-corrected time-domain traces for T211R1. (C) Overlap between predicted and experimental distance distributions for T211R1. The predicted distance distributions derived from solved crystal structures (TmPPase:Ca, TmPPase:Ca:ETD, TmPPase:ZLD, and TmPPase:IDP) are shown as dashed lines and overlaid with the corresponding DEER-derived distance distributions under each condition. Grey-shaded areas indicate the uncertainty associated with the DEER distributions. All data were processed using DeerAnalysis2022.

Overall, our DEER experiments highlight the asymmetric binding mode of TmPPase in solution and distinct inhibitor-specific effects on the conformational landscape, resulting in various closed-open configurations. IDP induces a closed-open state on both sides of the membrane, consistent with the presence of Na^+ in the ion gate[14], while ETD and ZLD generate a closed-open state on the cytoplasmic side, but an open-open states on the periplasmic side, consistent with no Na^+ in the ion gate (**Figure 17, Study I, Figure EV7**). Collectively, the combination of X-ray crystallography and solution-state DEER data allowed us to propose models outlining the conformational changes triggered by different compounds.

4.1.3 Conformational ensembles of TmPPase in solution

We proposed four asymmetric models to clearly illustrate the *apo*/-Ca/+Ca:ETD-state (model 1) and to highlight the distinct binding patterns of various inhibitors (ETD, ZLD and IDP; models 2-4) (**Figure 16**). Model 1 depicts an asymmetric state at the cytoplasmic side under *apo*, +Ca, and +Ca/ETD conditions. Loops5-6 are highly flexible, consistent with the broad distribution observed in DEER data for T211R1 and the absence of electron densities in the corresponding regions of the crystal structures. The periplasmic side remains in the 'open-open' state, with distance distribution largely unchanged from ~3.5 to 3.7nm, and

TMH 12 and TMH 16 positioned ‘up’, in agreement with the solved structures[47]. Model 2 describes the structural response to ETD binding. C599R1, located on TMH14, reports a ‘closed-open’ state, with ligand binding to just one active site. However, the modal distance distribution of ~ 3.9 nm observed at S525R1, along with the corresponding Bhattacharyya coefficient, indicates the absence of a complete conformational change on the periplasmic side, which remains in an ‘open-open’ state. Model 3, corresponding to ZLD binding, features the bulky heteroaryl group of ZLDs pushing the TMH 12 away (**Figure 11C**), resulting in a ‘partially closed – open’ state at the cytoplasmic side. In solution, this bulky group further hinders conformational changes on the periplasmic side, maintaining an ‘open–open’ state by preventing the full downward movement of TMH12 and TMH16[14,47]. Model 4, with IDP, induces a ‘closed-open’ state at the cytoplasmic side, again with ligand binding to just one active site. Importantly, IDP also drives a full downward movement of TMH 12 in one monomer, reflected in an increased distance distribution (~4.0 nm) at S525R1. This conformational shift results in an asymmetric conformation at the periplasmic side, while the other monomer remains open, consistent with the hybrid ‘closed-open’ structure of TmPPase:IDP(A)_Ca(B).

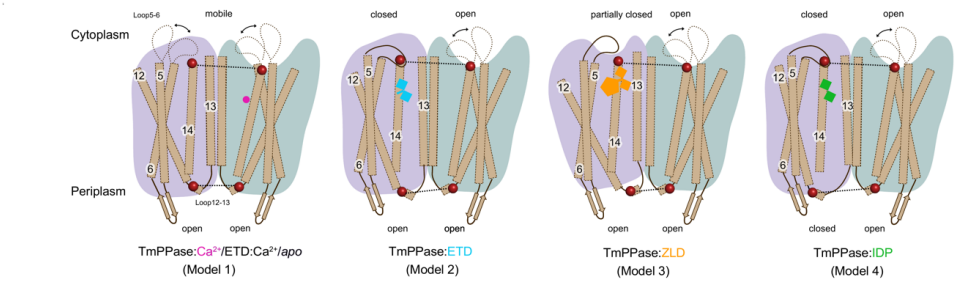


Figure 16. Models based on DEER distance distributions for TmPPase C599R1, S525R1, and T211R1. Four DEER models showing major conformational ensembles of TmPPase in solution. Two monomers are colored purple and green, respectively. All TMHs are shown in brown; mobile loop5-6 is indicated by a black dashed line, while fixed loop5-6 and loop12-13 are indicated by a solid black line; The labelling sites are represented by maroon spheres. Ca^{2+} is shown as a magenta circle; IDP is shown as purple squares; ETD as cyan squares connected by a cyan stick; ZLD as an orange pentagon.

To validate the presence of subtle conformational differences at the periplasmic side, we further assessed Na^+-pumping activity under ETD, ZLD, and IDP conditions. As previously discussed, the complete downward movement of TMH12 facilitates the formation of the ion gate required for Na^+ capture. Therefore, the membrane potential generated during pumping can serve as an indicator of whether only IDP induces a 'closed–open' conformation at the periplasmic side.

4.2 Pumping-after-hydrolysis mechanism

We performed electrometric measurements using solid-supported membrane-based electrophysiology[125] on TmPPase to investigate the energy coupling between PP_i hydrolysis and Na^+ pumping. These experiments aimed to address key questions regarding the sequence of ion transport and hydrolysis in mPPase, as well as to validate DEER models. This measurement quantified the total currents generated as Na^+ crossed the membrane through reconstituted TmPPase at RT. K_2HPO_4 was used as a negative control, and no significant signal was observed, indicating no ion-pumping event occurred (**Figure 17A**). A maximum positive signal of 0.6 ± 0.025 nA was detected following the addition of 100 µM substrate K_4PP_i. Replacing K_4PP_i with IDP resulted in a reduced current signal, approximately 50% of the K_4PP_i signal. Substituting IDP with ETD or ZLD further diminished the response, with signals becoming barely detectable.

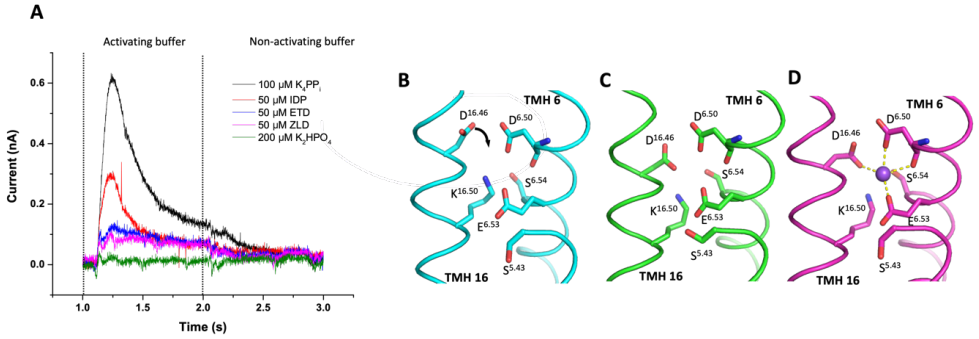


Figure 17. Transient currents of TmPPase Na^+ pumping and ion gate of TmPPase structures. (A) Curve of Na^+ pumping current triggered by 100 µM of K_4PP_i, 50 µM of IDP, 50 µM of ETD, 50 µM of ZLD, and 200 µM of K_2HPO_4. The vertical black dashed line represents the addition of activating buffer and non-activating buffer. (B-D). Ion gate of TmPPase:ETD (cyan); TmPPase:ZLD (green); TmPPase:IDP (purple). The black arrows show the movement of residues of $D^{16.46}$ and $K^{16.50}$.

In our previous study[14], we found that IDP induced membrane potential changes similar to PP_i; however, the total proton current signals in the presence of IDP was significantly lower than that observed with PP_i in VrPPase, where the IDP-induced signal was an order of magnitude lower than that of PP_i. There are three possible interpretations for this reduction, each supporting different mechanisms: 1. Ion pumping can occur without hydrolysis, but it cannot sustain multiple turnover events compared to PP_i, leading to reduced signals. 2. Partial ion translocation in the membrane, rather than a complete single turnover event, resulted in the weaker signal. 3. The signal originates from a mere reorientation of charged amino acids in functional regions, rather than ion translocation. The first interpretation was ruled out by the recent work of Baykov *et al.*,[62] whose stopped-flow pyranine assay demonstrated that proton pumping occurs exclusively in the presence of PP_i. This was indicated by the fluorescence decrease of the pH-sensitive dye pyranine under PP_i conditions. In contrast, the addition of IDP or bisphosphonates failed to produce a similar fluorescence change, suggesting no proton pumping occurred.

In our this Nanion SURFE2R experiment, the observed reduction in Na^+ current with IDP (compared to PP_i) can be attributed to the second mechanisms. This observation, combined with the TmPPase:IDP crystal structure, which reveals a Na^+ ion at the ion gate, suggests that IDP binding can induce Na^+ translocation to the ion gate, but its release requires PP_i. The weak current signals observed for ETD and ZLD, together with their bound structures, where no Na^+ was observed at the ion gate (**Figure 17B-C**), suggest that these signals likely arose from the reorientation of charged amino acids rather than actual ion translocation. Overall, our electrometric data are consistent with the findings of the stopped-flow pyranine assay by Baykov *et al*.,[62] further supporting the "pumping-after-hydrolysis" model, which posits that full ion pumping requires the substrate Mg_2PP_i. Moreover, our results demonstrate that subtle conformational changes at the periplasmic side are critical for enabling the complete downward movement of TMH 12—a structural rearrangement necessary for Na^+ ions to reach the ion gate. Accordingly, the 4.0 nm DEER distance observed at S525R1sites in the IDP-bound sample likely represents the minimal structural configuration required for efficient Na^+ translocation.

4.3 Characterization of mPPase from *Plasmodium falciparum*

So far, the structures of thermostable mPPases, including TmPPase, VrPPase, and PaPPase, as well as potential inhibitors against TmPPase, have been well characterized. In contrast, structural information on parasitic mPPase and their corresponding inhibitors remains limited. Moreover, translating these inhibitors into therapeutic applications requires testing on purified parasitic mPPase rather than relying solely on TmPPase as a model, due to pharmacological differences, as previously demonstrated[59]. In this project, we focus on the expression and purification of type 1 mPPase from *Plasmodium falciparum* (PfPPase), which is essential for ring-stage development and the transition to the trophozoite stage of the parasite[136] We first optimized an express system suitable for PfPPase, using an baculovirus/insect cell system, where Hi5 cells showed eightfold higher PfPPase expression levels than Sf9 cells (**Study III, Figure 1**), while other express systems failed. PfPPase was subsequently solubilized in a mixture of 1% DDM with 0.2% CHS within 6 h, achieving the highest yield of 88% (**Figure 18A-D**). Purification using IMAC yielded 0.3 mg per liter of PfPPase (**Figure 18E**).

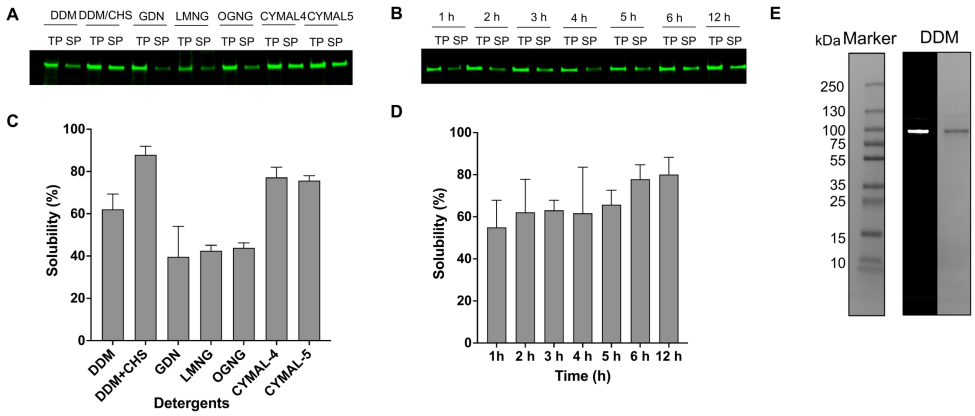


Figure 18. Detergent solubilization screening of Hi5 expressed His-tagged PfPPase fused with a C-terminal GFP and SDS-page of PfPPase (A) The in-gel fluorescence shows the solubilization efficiency of PfPPase from the following detergents: 1% DDM, 1% DDM + 0.2%CHS, 1% GDN, 1% LMNG, 1% OG, 1% CYMAL-4, and 1% CYMAL-5. (B) The in-gel fluorescence shows the solubilization efficiency of PfPPase over time. (C, D) The ratio of band intensities of solubilized protein (SP) to total protein (TP) was analyzed using Sapphire FL and solubility percentage was calculated using ImageJ. (E) The GFP fluorescence and the Coomassie stained SDS-Page showing the purified PfPPase in DDM.

4.3.1 The effect of detergents and C-terminal fusion on the activity of PfPPase

Although DDM is the most effective detergent for solubilizing PfPPase, GDN and LMNG better preserved its enzymatic activity better (142 ± 7 nmol·μg^{-1}·min^{-1}; 120 ± 3 nmol·μg^{1}·min^{-1}) compared to DDM (103 ± 2 nmol·μg^{-1}·min^{-1}) ($p < 0.0001$) (**Figure 19**). This improved activity retention is likely due to the stronger detergent-detergent interactions in GDN[137], reducing protein aggregation. Similarly, LMNG, composed of two DDM-like molecules, provides greater hydrophobic surface coverage for membrane proteins, enhancing `stability[137]. Therefore, exchanging DDM for GDN or LMNG during purification improves downstream applications.

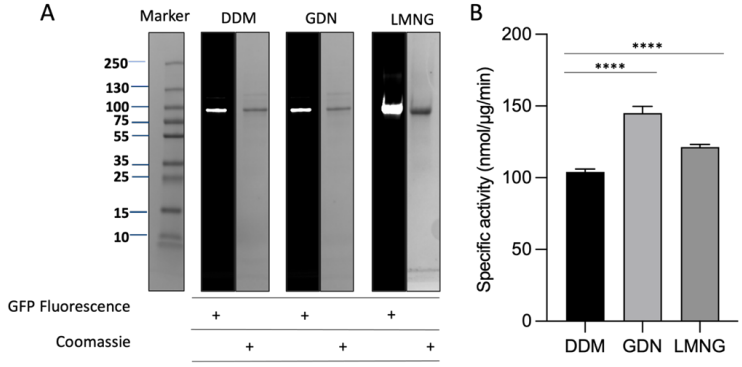


Figure 19. In-gel fluorescence and Coomassie-stained SDS-page of PfPPase. (A) The GFP fluorescence and the Coomassie stained SDS-Page showing the purified PfPPase in DDM, GDN, and LMNG. (B) The bar

chart showing the specific activity of PfPPase in DDM (103 ± 2 nmol·μg^{-1}·min^{-1}), GDN (142 ± 7 nmol·μg^{-1}·min^{-1}) and LMNG (120 ± 3 nmol·μg^{-1}·min^{-1}) ($p < 0.0001$).

To monitor PfPPase expression, a GFP tag was fused to its C-terminus[138]. However, the GFP tag could not be cleaved (**Study III, Figure S1**). To address this, we designed, expressed, and purified a GFP-free construct using the same methods as those used for the GFP-tagged PfPPase (**Study III, Figure 5**). Activity assays revealed that the GFP-free construct exhibited threefold higher activity (520 ± 13 nmol·μg^{-1}·min^{-1}) than GFP-tagged PfPPase. This increase suggests that the C-terminal GFP tag may interfere with loop and helix movement near the exit channel disrupting the half-of-the-sites reactivity[48]. This mechanism relies on intra-subunit communication, particularly, a positive charge at $A^{12.46}$ (in K^+-dependent mPPase) or $K^{12.46}$ (in K^+-independent mPPase)[48]. In the presence of K^+, the downward motion of TMH12 in monomer A is necessary for triggering conformational changes at the dimer interface, mediated by exit-channel loop12-13 moving toward the dimer interface to activate hydrolysis in monomer B[59]. If this process is disrupted. For instance, the allosteric inhibitor ATC binds to the TmPPase exit channel *via* loop6-7, loop8-9, and loop12–13, effectively locking intra-subunit communication and reducing activity [59]. Additionally, the conserved total length of exit channel loop8-9 and loop10-11 suggests that they may also contribute to enzymatic activity through an as-yet-unknown mechanism[48]. Overall, the reduced activity of PfPPase with a C-terminally fused GFP further highlights the importance of conformational rearrangements in the exit channel during hydrolysis.

4.3.2 Determination of oligomer state of PfPPase

All mPPases characterized so far function as homodimers. To investigate whether PfPPase is also active in a dimeric state, we performed an in-gel CN activity assay. Coomassie-stained CN PAGE revealed two distinct bands for PfPPase, whereas TmPPase appeared as a single band (**Figure 20A**). Subsequently, the in-gel molybdenum reaction confirmed that enzymatic activity was exclusively with the upper band observed on the Coomassie-stained CN PAGE (**Figure 20B**), indicating that only the dimeric form is active, while the monomeric form is inactive. Further analysis using SEC-MALS and BN PAGE confirmed the presence of monomeric PfPPase (**Study III, Figure S3**), with a MW of 176 kDa, lower than the expected 206 kDa, suggesting a small proportion of monomers. These monomers comprised 8–26% of total PfPPase across three expression batches. Nevertheless, the exclusive activity of the dimeric form further supports the half-of-the-sites activity mechanism[48], where the mPPase dimer functions as an integrated machine. Conformational changes at the dimer interface, driven by the downward motion of TMH12, the reorganization of the exit-channel loops, and the upward motion of TMH13, are essential for catalytic activity[48,59].

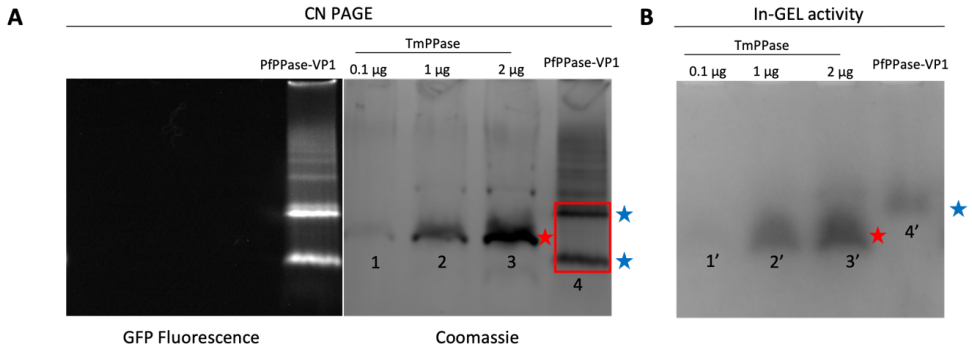


Figure 20. Fig 6. Clear Native PAGE and In-gel activity assay of PfPPase and TmPPase. (A) In-gel fluorescence and Coomassie stained CN PAGE of PfPPase and TmPPase showing two oligomeric states of PfPPase (blue stars) and the dimeric TmPPase (red star). (B) In-gel activity assay showing the black-colored complex formed by phosphate with molybdate, corresponding to the position of the upper band of PfPPase and TmPPase in (A).

4.4 Identification of inhibitors against PfPPase

4.4.1 Non-hydrolyzable PP_i analog inhibitors

In study I, we obtained two TmPPase crystal structures in complex with ETD and ZLD, which are non-hydrolyzable PP_i analogs. Additionally, we obtained more TmPPase crystals in complex with five other bisphosphonates; however, these crystals diffracted weakly. Activity assay revealed that among these compounds, ALE, PAM, and neridronate (NRD) exhibited poor inhibitory activity, with IC_{50} values ranging 100 and 200 µM (**Study I, Figure 1B**), while ETD and ZLD were inactive against TmPPase. The differences in inhibitory activity may be due to the introduction of an amino group in the side chain and the length of the side chain[139]. Substituting the hydrogen (in ETD) with a benzene ring (in ZLD and RSD) does not favour inhibitory activity (IC_{50} > 200 µM). Next, I tested these seven bisphosphonates and IDP on PfPPase, and none demonstrated inhibition except for IDP, which has an estimated IC_{50} of 182 [151-239] µM (**Figure 21A**).

4.4.2 Non-phosphorus inhibitors

We then evaluated seven previously identified TmPPase inhibitors and fourteen compounds with unknown activity on TmPPase for their potential to inhibit PfPPase (**Study III, Table 2 and S1**). Among these, only two TmPPase inhibitors derived from the pyrazolo[1,5-*a*]pyrimidine core inhibited PfPPase, with IC_{50} values below 100 µM (74 and 58 µM, respectively) (**Figure 21B-D**). However, these values indicate 4- to 10- fold lower inhibitory activity compared to their effects on TmPPase (**Study III, Table 2**). Additionally, we assessed the specificity of fourteen commercially available drugs targeting malaria parasites

on both PfPPase and TmPPase (**Study III, Table 1**). Of these, only furamidine (RD2) demonstrated weak activity, with an IC_{50} value of 180 [161-207] µM against PfPPase and no detectable activity against TmPPase (**Figure 21E**), while the other compounds had no activity against either PfPPase or TmPPase.

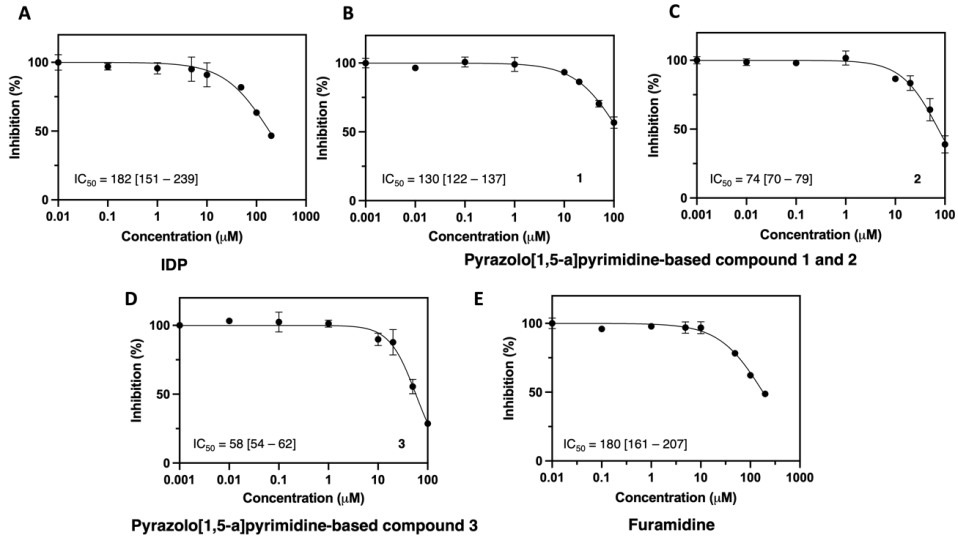


Figure 21. IC_{50} values of PfPPase inhibitors. (A-E). Inhibition of PfPPase by IDP (A), pyrazolo[1,5-*a*]pyrimidine-based compounds (B-D) and furamidine (E).

The pharmacological differences in the activity of pyrazolo[1,5-*a*]pyrimidine-based compounds between TmPPase and PfPPase can be attributed to several factors. For example, although they may share a binding site at the exit channel; however, these sites are not highly conserved, as evidenced by the absence of an ATC binding site in PfPPase[59]. Alternatively, differences in the accessibility of the conserved active site between the two mPPases, particularly if the volume of the hydrolytic center differs, may result in significant changes in binding affinity. Therefore, further characterization of the binding sites in TmPPase and PfPPase is essential to elucidate the structural basis underlying the observed differences in binding affinity between these enzymes. Such insights could facilitate the rational design of more selective and potent PfPPase inhibitors, with potential therapeutic implications for malaria treatment.

4.5 Preliminary structural studies of PfPPase

4.5.1 Complex reconstruction

PfPPase-BRIL_C was solubilized in DDM+CHS and purified in GDN, as previously described, due to the higher activity of PfPPase in GDN compared to DDM (section 4.3.1).

Moreover, GDN has emerged as a widely favored detergent for structural studies of membrane proteins, as it provides a more native-like lipid environment and statilizes dynamic or fragile complexes[140-142]. Notably, a number of high-resolution structures of membrane proteins purified in GDN have been reported[140-142].

PfPPase-BRIL and BAK5 were mixed and purified by SEC, as described in Section 3.1. The sample eluted as three peaks (**Figure 22A**), with the central peak being subjected to SDS-PAGE analysis (**Figure 22B**). The SDS-PAGE results demonstrated that the first peak fractions contained both PfPPase-BRIL and BAK5. The second peak fractions are believed to contain contaminants (~10 kDa), which was also observed by Nikolova[143]. However, the 10% gel was unable to resolve the overlapping bands. To address this, I used a 4–20% precast gel, which successfully separated the approximately 10-15 kDa contaminant (**Figure 22C**). Additionally, BAK5 appeared as two distinct bands that were not visible on the 10% gel. These two bands are attributed to the heavy (27 kDa) and light chains (26 kDa) of the antibody fragment[132]. The pooled fractions from the first peaks displayed a major band on BN-PAGE, corresponding to the homogeneous PfPPase-BRIL_C/BAK5 complex with an estimated MW ≥ 250 kDa (theoretical MW: 292 kDa). The adjacent band corresponds to PfPPase-BRIL_C, migrating between 130 and 250 kDa (theoretical MW: 186 kDa) (**Figure 22D**).

The homogeneity of the sample was further evaluated using negative-stain EM to determine whether it was of sufficient quality for cryo-EM analysis. Particles were evenly distributed on the grid and manually picked using a 200 Å mask, followed by extraction into a 250-pixel box. A total of 1,452 particles were extracted and subjected to 2D classification. The representative populated classes exhibited features consistent with expectations: a smaller blob (BAK5) attached to a larger blob, likely corresponding to the PfPPase-BRIL_C complex reconstituted in GDN micelles (**Figure 22F**). In contrast, the negative-stain EM **of** PfPPase alone revealed no distinctly populated classes (**Figure 22E**). Instead, individual PfPPase particles appeared embedded within detergent micelles, lacking discernible structural features. This poses a significant challenge for image processing and alignment, as the absence of well-defined boundaries prevents accurate classification and averaging during 2D analysis[144,145]. Beyond BRIL, a relevant example is the *Plasmodium falciparum* chloroquine resistance transporter (PfCRT), a 49 kDa integral membrane protein involved in antimalarial drug resistance[146]. Using a phage display library, Kim *et al*. identified high-affinity Fabs, which enhanced particle alignment and facilitated the determination of the PfCRT structure at 3.2 Å resolution, revealing its overall architecture and a potential drug-binding site[146]. Overall, the complex can be observed in negative stain EM and the particles were well-dispersed after SEC. The sample was taken forward for cryo-EM grid screening.

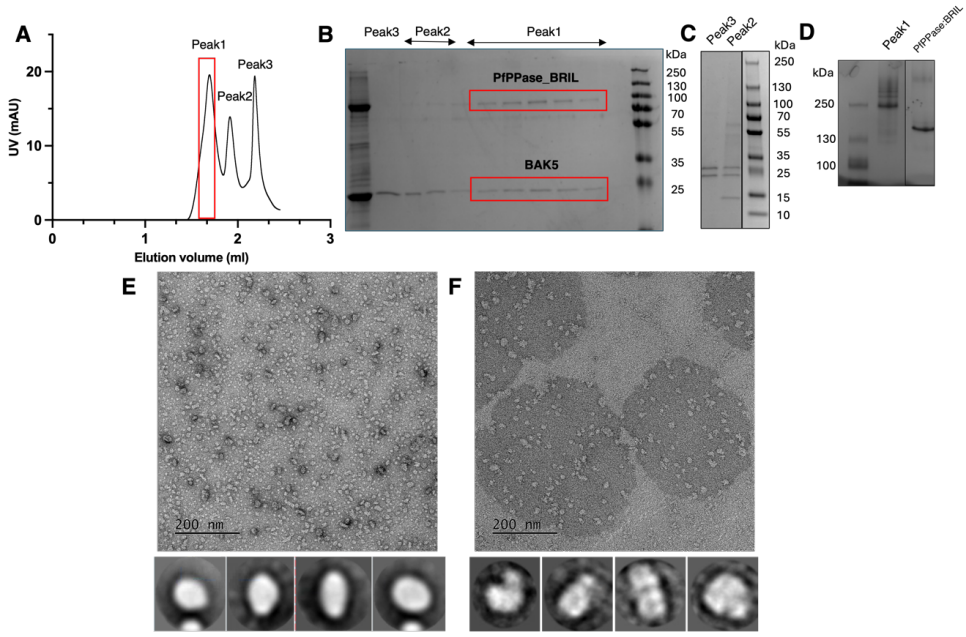


Figure 22. The formation of PfPPase-BRIL_C/BAK5 complex. (A). SEC profile showing three major peak fractions from the purification of complex. These fractions were collected separately for SDS-PAGE analysis. The peak fraction of complex is highlighted with red rectangles. (B). SDS-PAGE analysis of PfPPase-BRIL_C/BAK5 complex from the SEC peak fractions using a 12%-page gel under reducing conditions. (C). SDS-PAGE analysis of peak 1and peak 2 fractions from the SEC using 4 -20% page gel under reducing conditions. (D) BN-PAGE shows the homogeneous complex from the peak 1 fraction and PfPPase-Bril_C without BAK5. (E). An example of a negative stain micrograph of purified PfPPase and representative 2D classes. (F) An example of a negative stain micrograph of purified PfPPase-BRIL_C/BAK5 complex and representative 2D classes. The images were collected on a Hitachi microscope, and 2D classification was performed using 3.8 Scipion.

4.5.2 Cryo-EM grid optimization

The purified PfPPase-BRIL_C/BAK5 complex was applied to various types of grids for initial trials. Despite the good distribution of homogeneous particles observed on negative-stain EM grids, particles were scarcely found on Quantifoil Cu and Au-flat grids (**Figure 23A-B**). The "particles" observed on the C-flat Au grid could not be aligned (**Figure 23C**) and are thus presumed to be artifacts generated during vitrification or grid transfer. This suggested that the protein aggregated on the support film rather than occupies the holes[147]. To address this, grids with a 2 nm continuous carbon film or graphene oxide coating on the support film were used to concentrate the protein within the holes[147]. Using these grids, protein particles became visible; however, the complex particles tended to aggregate on both types of grids (**Figure 23D-E**). Aggregation likely occurred during the vitrification process, possibly due to interactions at the air-water interface (AWI)[148], or the high protein concentration used (1 mg/ml) compared to the lower concentration (0.01 mg/ml) employed in negative-stain EM. To mitigate aggregation, 3 mM Fos-Choline-8 was added to the

sample immediately prior to vitrification, a strategy previously reported to reduce aggregation and improve protein distribution on cryo-EM grids[108,149,150]. This addition significantly improved the distribution of protein particles at a sample concentration of 2.5 mg/ml (**Figure 23F**).

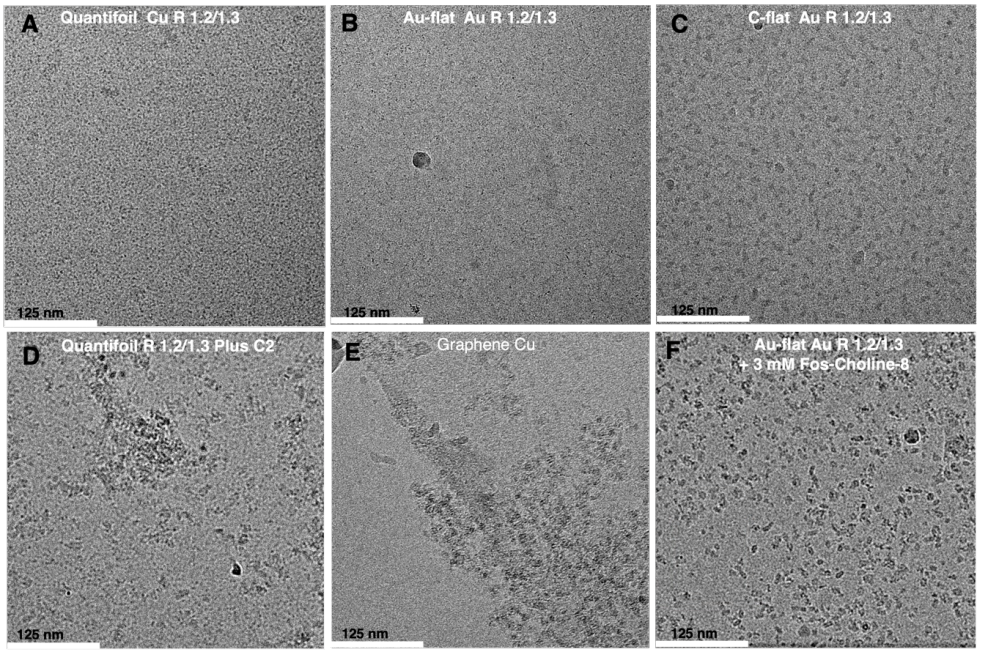


Figure 23. Cryo-EM grid optimization. (A-F). Cryo-EM images show the particle distribution of the PfPPase-BRIL_C/BAK5 complex sample using different grid supports and additives. Images were collected on a Titan Krios microscope at 150,000x magnification.

Following specimen optimization, a large dataset was collected using the Au-flat Au R1 1.2/1.3 grid (**Figure 23F**) for subsequent analysis. After multiple rounds of 2D classification, 184,479 particles were selected for the final 2D classification (**Figure 24**). However, the majority of classes lacked recognizable features and could not be identified. These particles were likely GDN micelles (personal communication) and/or protein that had dissociated from BAK5. Notably, in the seventh and tenth classes of the second row, as well as the fifth and seventh classes of the third row, two small densities ("blobs") attached to the micelle were observed, potentially corresponding to two BAK5 molecules bound to BRIL in each monomer of the PfPPase-BRIL_C. Unfortunately, due to the prevalence of low-quality particles, the 3D initial models obtained from the dataset were not interpretable.

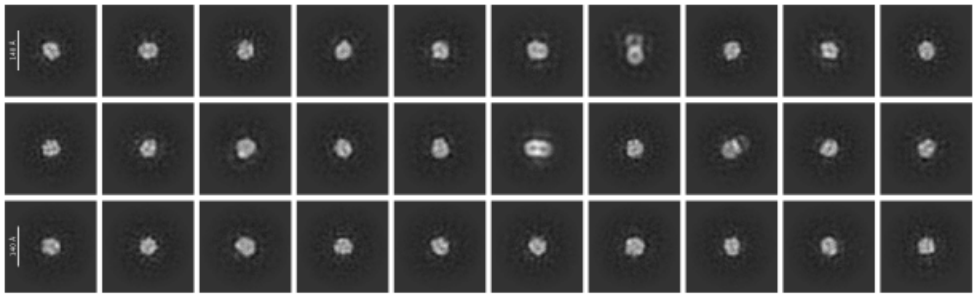


Figure 24. Initial 2D classes obtained from the large dataset of the purified PfPPase-BRIL/BAK5 complex. Particles were extracted in a 500-pixel box. The classes are arranged based on the number of particles contributing to each, starting with the class containing the highest number of particles at the top left and descending to the bottom right.

Determining the structure of PfPPase-BRIL_C by cryo-EM presents two major challenges: protein instability under cryo-EM conditions and interference from detergents**.** These factors significantly affect particle integrity, orientation control, and resolution. A promising alternative for future studies is PfPPase-BRIL_N, which exhibits higher enzymatic activity compared to PfPPase-BRIL_C. As discussed above, the C-terminal BRIL fusion may interact with the exit channel loops, affecting enzymatic activity and potentially destabilizing the protein by disrupting the dimer interface. Consequently, this construct failed to achieve its intended goal of providing structural insights into the role of the exit channel in catalysis. Therefore, PfPPase-BRIL_N may offer improved structural stability, making it more suitable for cryo-EM analysis.

To overcome the limitations of detergents, detergent-free approaches such as nanodiscs and styrene-maleic anhydride (SMA) polymers have been explored[151]. Nanodiscs comprises a lipid bilayer stabilized by membrane scaffold proteins, providing a native-like environment for membrane proteins. However, attempts to reconstitute PfPPase into nanodiscs were unsuccessful, as the protein failed to incorporate effectively. Alternatively, Styrene-Maleic Anhydride (SMA) copolymers can extract membrane proteins directly from lipid bilayers, forming SMA-lipid particles (SMALPs)[152]. Unfortunately, PfPPase exhibited aggregation in the presence of SMA, indicating incompatibility with this system. **In Study III**, LMNG was shown to retain enzymatic activity comparable to GDN, even at concentrations below its critical micelle concentration (CMC) of 0.001% w/w (**Figure 19A**). One key advantage of LMNG is its low CMC, which results in a slow off-rate, allowing it to remain bound to the protein surface even at very low concentrations[153]. This property has been demonstrated in the β_2-adrenergic receptor, where LMNG-stabilized receptors retained their ability to bind G-proteins at detergent concentrations 1,000-fold below the CMC[154]. Therefore, LMNG will be interesting for future cryo-EM studies to minimize background noise caused by detergent micelles.

5 Conclusion and Future Studies

In summary, this thesis has provided detailed mechanistic insights into mPPase biochemistry, conclusively demonstrating its asymmetric catalytic mode *via* DEER spectroscopy and crystallography. The successful production of PfPPase as a model enzyme, combined with ongoing cryo-EM studies, paves the way for structure-guided drug discovery targeting parasitic mPPases.

Studies I and **II** demonstrated that solution-state EPR experiments, combined with newly solved structures of TmPPase bound to bisphosphonates inhibitors, confirm asymmetric inhibitor binding and the formation of multiple conformational states under different inhibitory conditions. These findings provide strong evidence for symmetry-breaking across the membrane, consistent with concept of half-of-the-sites-reactivity[48]. To our knowledge, **Study I** presents the first to measurements of conformational changes in mPPase that mimic the substrate hydrolysis process under physiological conditions. Solid-supported membrane-based electrophysiology recordings in **study I** and **II** were used to investigate the temporal relationship between Na^+ pumping and PP_i hydrolysis. The results indicate that Na^+ pumping is tightly coupled to PP_i hydrolysis. In the absence of hydrolysis, Na^+ pumping is restricted to the ion gate in the presence of IDP, consistent with the TmPPase:IDP structure, where Na^+ release from the ion gate is not possible. Furthermore, the barely detectable Na^+ signals in the presence of ETD and ZLD suggests that one monomer must adopt a fully closed state on both side of membrane for Na^+ pumping or partial translocation to occur.

To identify non-phosphorus inhibitors targeting mPPase, it is crucial to validate their pharmacological activity using parasitic mPPases. In **studies III** and **IV**, I established an approach encompassing the expression and purification of PfPPase, along with the exploration of potential novel inhibitors. The Hi5 insect cell expression system proved highly efficient for expressing various PfPPase constructs, thereby facilitating future investigation into its mechanism and structure. I identified two pyrazolo[1,5-*a*]pyrimidine-based compounds with low micromolar activity (<100 μM) against PfPPase. However, the inhibitors developed to date have primarily been designed based on the TmPPase model and exhibit greater specificity for TmPPase than for PfPPase. Unpublished structural data presented in this thesis demonstrate optimizations in cryo-EM grid preparation, providing a

viable pathway to resolve the PfPPase structure—an essential step for the structure-based design of more selective inhibitors.

Looking ahead, cryo-trapped time-resolved DEER could be employed in conjunction with substrate-bound complexes to investigate, for example, the order and mechanism by which the release of phosphate products and ions induces conformational changes in the active sites, exit channels, and inter-subunit communication. This approach could provide deeper insights into the reaction steps and their structural dynamics. Additionally, expanded screening efforts using diverse scaffolds, including azulene-based series, will be carried out against both TmPPase and PfPPase to identify new inhibitory compounds. Ultimately, resolving the structure of PfPPase remains a critical milestone for the rational design of selective and potent inhibitors.

6 Acknowledgements

The work for described in this thesis was carried out between 2020 and 2024 at the Molecular and Integrative Biosciences, University of Helsinki. It was financed by the China Scholarship Council (CSC) from the Ministry of Education of P.R. China, Biotechnology and Biological Research Council (BBSRC), and Academy of Finland.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Adrian Goldman, for his invaluable guidance, support, and encouragement throughout my Ph.D. journey. His expertise, patience, and insightful feedback have been instrumental in shaping both this thesis and my growth as a researcher. And my co-supervisor, Dr Henri Xhaard, showed genuine care and encouragement for both my research and writing progress, helping to ensure I could graduate smoothly. Despite his busy schedule, he always made time for regular meetings to discuss my work, for which I am deeply grateful.

I would also like to thank the members of my thesis committee for their time, Prof. Kari Keinanen and Dr. Tommi Kajander. Thanks for your constructive comments, and support during each stage of this work. I thank Prof. Lars Jeuken for warmly welcoming me on my visits to Netherland.

Special thanks to co-authors and my collaborators in the University of Helsinki and beyond: Dr. Anokhi Shah, Xinyu Liu, Dr. Joshua L. Wort, Dr. Yue Ma, Dr. Katie Hardman, Dr. Niklas G. Johansson, Dr. Orquidea Ribeiro1, Dr. Adam Brookfield, Dr. Alice Bowen, Prof. Jari Yli-Kauhaluoma, Dr. Keni Vidilaseris, and Assistant Prof. Christos Pliotas. Your advice, discussions, and shared experiences have enriched this project in many ways.

It has been my great pleasure to work with all the members of the Goldman group. Keni, I am deeply grateful for your unwavering support and guidance throughout my work. You patiently led me step by step, and I will never forget the long nights we spent revising manuscripts together. Orquidea, I am sincerely grateful for your support in helping me integrate into the lab's daily routines and experimental work. Your excellent technical skills greatly eased my project and helped me avoid many setbacks. I thank my lab mates, Sadegh, Mimmu, and Paulina, for your support during experiments and your engaging and constructive discussions in group meetings, which helped advance my research. I also thank

our technician, Tukka, for your thoughtful care of everything in the lab, and for always being the first person I could turn to whenever I needed help.

To my friends: Zhehao, Ruoxi, Yuwen, Danni, and Jack. I will always cherish the wonderful memories of our travels across many countries and the joyful weekend gatherings. It was you who helped me find joy in life amid the busyness of lab work.

最后，我想用中文感谢我的父母，是你们的鼓励让我走到了今天，没有人知道我有多爱你们，只希望未来我们可以永远在一起，还有很多家庭时光一起度过，祝你们身体健康，永远快乐。

7 References

1. Lahti, R. Microbial inorganic pyrophosphatases. *Microbiol Rev* **47**, 169–178 (1983).
2. Terkeltaub, R. A. Inorganic pyrophosphate generation and disposition in pathophysiology. *Am J Physiol Cell Physiol* **281**, C1–C11 (2001).
3. Farquharson, K. L. Life of PPi: soluble PPases and H^+-PPase act cooperatively to keep pyrophosphate levels in Check. *Plant Cell* **30**, 951–951 (2018).
4. Yang, X. *et al*. Essential role of pyrophosphate homeostasis mediated by the pyrophosphate-dependent phosphofructokinase in *Toxoplasma gondii*. *PLoS Pathog* **18**, e1010293 (2022).
5. Kajander, T., Kellosalo, J. & Goldman, A. Inorganic pyrophosphatases: one substrate, three mechanisms. *FEBS Lett* **587**, 1863–1869 (2013).
6. Cooperman, B. S., Baykov, A. A. & Lahti, R. Evolutionary conservation of the active site of soluble inorganic pyrophosphatase. *Trends Biochem Sci* **17**, 262–266 (1992).
7. Josse, J. & Wong, S. C. K. Inorganic pyrophosphatase of *Escherichia coli* in *The enzymes* Vol. 4 (ed. Boyer, P. D.) 499–527 (Academic Press, 1971).
8. Bunick, G., McKenna, G. P., Colton, R. & Voet, D. The X-ray structure of yeast inorganic pyrophosphatase: CRYSTAL PROPERTIES. *J Biol Chem* **249**, 4647–4649 (1974).
9. Heikinheimo, P. *et al*. The structural basis for pyrophosphatase catalysis. *Structure* **4**, 1491–1508 (1996).
10. Samygina, V. R. Inorganic pyrophosphatases: structural diversity serving the function. *Russ Chem Rev* **85**, 464 (2016).
11. Merckel, M. C. *et al*. Crystal structure of *Streptococcus mutans* pyrophosphatase: a new fold for an old mechanism. *Structure* **9**, 289–297 (2001).
12. Fabrichniy, I. P. *et al*. A trimetal site and substrate distortion in a family II inorganic pyrophosphatase. *J Biol Chem* **282**, 1422–1431 (2007).
13. Baltscheffsky, H., Von Stedingk, L. V., Heldt, H. W. & Klingenberg, M. Inorganic pyrophosphate: formation in bacterial photophosphorylation. *Science* **153**, 1120–1122 (1966).
14. Li, K. M. *et al*. Membrane pyrophosphatases from *Thermotoga maritima* and *Vigna radiata* suggest a conserved coupling mechanism. *Nat Commun* **7**, 13596 (2016).
15. Lin, S. M. *et al*. Crystal structure of a membrane-embedded H^+-translocating pyrophosphatase. *Nature* **484**, 399–403 (2012).
16. Moyle, J., Mitchell, R. & Mitchell, P. Proton-translocating pyrophosphatase of *Rhodospirillum rubrum*. *FEBS Lett* **23**, 233–236 (1972).
17. Malinen, A. M., Belogurov, G. A., Baykov, A. A. & Lahti, R. Na^+-pyrophosphatase: a novel primary sodium pump. *Biochemistry* **46**, 8872–8878 (2007).
18. Luoto, H. H., Baykov, A. A., Lahti, R. & Malinen, A. M. Membrane-integral pyrophosphatase subfamily capable of translocating both Na^+ and H^+. *Proc Natl Acad Sci USA* **110**, 1255–1260 (2013).
19. Baltscheffsky, M., Schultz, A. & Baltscheffsky, H. H^+-PPases: a tightly membrane-bound family. *FEBS Lett* **457**, 527–533 (1999).
20. Baykov, A. A., Malinen, A. M., Luoto, H. H. & Lahti, R. Pyrophosphate-fueled Na^+ and H^+ transport in prokaryotes. *Microbiol Mol Biol Rev* **77**, 267–276 (2013).
21. Walker, R. R. & Leigh, R. A. Mg^{2+}-dependent, cation-stimulated inorganic pyrophosphatase associated with vacuoles isolated from storage roots of red beet (*Beta vulgaris* L.). *Planta* **153**, 150–155 (1981).

22. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors in *Methods in neurosciences* Vol. 25 (ed. Sealfon, S. C.) 366–428 (Academic Press, 1995).
23. Nordbo, E., Luoto, H. H., Baykov, A. A., Lahti, R. & Malinen, A. M. Two independent evolutionary routes to Na^+/H^+ cotransport function in membrane pyrophosphatases. *Biochem J* **473**, 3099–3111 (2016).
24. Luoto, H. H., Nordbo, E., Malinen, A. M., Baykov, A. A. & Lahti, R. Evolutionarily divergent, Na^+-regulated H^+-transporting membrane-bound pyrophosphatases. *Biochem J* **467**, 281–291 (2015).
25. Yoon, H. S., Kim, S. Y. & Kim, I. S. Stress response of plant H^+-PPase-expressing transgenic *Escherichia coli* and *Saccharomyces cerevisiae*: a potentially useful mechanism for the development of stress-tolerant organisms. *J Appl Genet* **54**, 129–133 (2013).
26. Luoto, H. H., Belogurov, G. A., Baykov, A. A., Lahti, R. & Malinen, A. M. Na^+-translocating membrane pyrophosphatases are widespread in the microbial world and evolutionarily precede H^+-translocating pyrophosphatases. *J Biol Chem* **286**, 21633–21642 (2011).
27. Huber, R. *et al*. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Arch Microbiol* **144**, 324–333 (1986).
28. Huber, R. *et al*. *Pyrobaculum aerophilum* sp. nov., a novel nitrate-reducing hyperthermophilic archaeum. *Appl Environ Microbiol* **59**, 2918–2926 (1993).
29. Kellosalo, J., Kajander, T., Palmgren, M., Lopéz-Marqués, R. L. & Goldman, A. Heterologous expression and purification of membrane-bound pyrophosphatases. *Protein Expr Purif* **79**, 25–34 (2011).
30. Kriegel, A. *et al*. Job sharing in the endomembrane system: vacuolar acidification requires the combined activity of V-ATPase and V-PPase. *Plant Cell* **27**, 3383–3396 (2015).
31. Maeshima, M. Vacuolar H^+-pyrophosphatase. *Biochim Biophys Acta Biomembr* **1465**, 37–51 (2000).
32. Li, J. *et al*. Arabidopsis H^+-PPase AVP1 regulates auxin-mediated organ development. *Science* **310**, 121–125 (2005).
33. Ferjani, A. *et al*. Keep an eye on PPi: the vacuolar-type H^+-pyrophosphatase regulates postgerminative development in *Arabidopsis*. *Plant Cell* **23**, 2895–2908 (2011).
34. Li, J. *et al*. *Arabidopsis* H^+-PPase AVP1 regulates auxin-mediated organ development. *Science* **310**, 121–125 (2005).
35. Park, S. *et al*. Up-regulation of a H^+-pyrophosphatase (H^+-PPase) as a strategy to engineer drought-resistant crop plants. *Proc Natl Acad Sci USA* **102**, 18830–18835 (2005).
36. Brini, F., Hanin, M., Mezghani, I., Berkowitz, G. A. & Masmoudi, K. Overexpression of wheat Na^+/H^+ antiporter TNHX1 and H^+-pyrophosphatase TVP1 improve salt-and drought-stress tolerance in *Arabidopsis thaliana* plants. *J Exp Bot* **58**, 301–308 (2007).
37. Zhao, F. Y., Zhang, X. J., Li, P. H., Zhao, Y. X. & Zhang, H. Co-expression of the *Suaeda salsa* SsNHX1 and *Arabidopsis* AVP1 confer greater salt tolerance to transgenic rice than the single SsNHX1. *Mol Breed* **17**, 341–353 (2006).
38. Zhang, H., Shen, G., Kuppu, S., Gaxiola, R. & Payton, P. Creating drought-and salt-tolerant cotton by overexpressing a vacuolar pyrophosphatase gene. *Plant Signal Behav* **6**, 861–863 (2011).
39. Pasapula, V. *et al*. Expression of an *Arabidopsis* vacuolar H^+-pyrophosphatase gene (AVP1) in cotton improves drought-and salt tolerance and increases fibre yield in the field conditions. *Plant Biotechnol J* **9**, 88–99 (2011).
40. Crompton, P. D. *et al*. Malaria immunity in man and mosquito: insights into unsolved mysteries of a deadly infectious disease. *Annu Rev Immunol* **32**, 157–187 (2014).
41. Docampo, R. & Moreno, S. N. Acidocalcisomes. *Cell Calcium* **50**, 113–119 (2011).
42. Lander, N., Cordeiro, C., Huang, G. & Docampo, R. Polyphosphate and acidocalcisomes. *Biochem Soc Trans* **44**, 1–6 (2016).

43. Lefurgey, A., Ingram, P. & Blum, J. J. Compartmental responses to acute osmotic stress in *Leishmania major* result in rapid loss of Na^+ and Cl^-. *Comp Biochem Physiol A Mol Integr Physiol* **128**, 385–394 (2001).
44. Lemercier, G. *et al*. A vacuolar-type H^+-pyrophosphatase governs maintenance of functional acidocalcisomes and growth of the insect and mammalian forms of *Trypanosoma brucei*. *J Biol Chem* **277**, 37369–37376 (2002).
45. Liu, J. *et al*. A vacuolar-H^+-pyrophosphatase (TgVP 1) is required for microneme secretion, host cell invasion, and extracellular survival of *Toxoplasma gondii*. *Mol Microbiol* **93**, 698–712 (2014).
46. Zhang, M. *et al*. Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* **360**, eaap7847 (2018).
47. Kellosalo, J., Kajander, T., Kogan, K., Pokharel, K. & Goldman, A. The structure and catalytic cycle of a sodium-pumping pyrophosphatase. *Science* **337**, 473–476 (2012).
48. Strauss, J. *et al*. Functional and structural asymmetry suggest a unifying principle for catalysis in membrane-bound pyrophosphatases. *EMBO Rep* **25**, 853–875 (2024).
49. Nakanishi, Y., Saijo, T., Wada, Y. & Maeshima, M. Mutagenic analysis of functional residues in putative substrate-binding site and acidic domains of vacuolar H^+-pyrophosphatase. *J Biol Chem* **276**, 7654–7660 (2001).
50. Schultz, A. & Baltscheffsky, M. Properties of mutated *Rhodospirillum rubrum* H^+-pyrophosphatase expressed in *Escherichia coli*. *Biochim Biophys Acta Bioenerg* **1607**, 141–151 (2003).
51. Hirono, M., Nakanishi, Y. & Maeshima, M. Essential amino acid residues in the central transmembrane domains and loops for energy coupling of *Streptomyces coelicolor* A3(2) H^+-pyrophosphatase. *Biochim Biophys Acta Bioenerg* **1767**, 930–939 (2007).
52. Asaoka, M., Segami, S. & Maeshima, M. Identification of the critical residues for the function of vacuolar H^+-pyrophosphatase by mutational analysis based on the 3D structure. *J Biochem* **156**, 333–344 (2014).
53. Tsai, J. Y. *et al*. Roles of the hydrophobic gate and exit channel in *Vigna radiata* pyrophosphatase ion translocation. *J Mol Biol* **431**, 1619–1632 (2019).
54. Wu, J. J., Ma, J. T. & Pan, R. L. Functional size analysis of pyrophosphatase from *Rhodospirillum rubrum* determined by radiation inactivation. *FEBS Lett* **283**, 57–60 (1991).
55. Sarafian, V., Potier, M. & Poole, R. J. Radiation-inactivation analysis of vacuolar H^+-ATPase and H^+-pyrophosphatase from *Beta vulgaris* L. Functional sizes for substrate hydrolysis and for H^+ transport. *Biochem J* **283**, 493–497 (1992).
56. Tzeng, C. M. *et al*. Subunit structure of vacuolar proton-pyrophosphatase as determined by radiation inactivation. *Biochem J* **316**, 143–147 (1996).
57. Malinen, A. M. *et al*. Pre-steady-state kinetics and solvent isotope effects support the "billiard-type" transport mechanism in Na^+-translocating pyrophosphatase. *Protein Sci* **31**, e4394 (2022).
58. Artukka, E., Luoto, H. H., Baykov, A. A., Lahti, R. & Malinen, A. M. Role of the potassium/lysine cationic center in catalysis and functional asymmetry in membrane-bound pyrophosphatases. *Biochem J* **475**, 1141–1158 (2018).
59. Vidilaseris, K. *et al*. Asymmetry in catalysis by *Thermotoga maritima* membrane-bound pyrophosphatase demonstrated by a nonphosphorus allosteric inhibitor. *Sci Adv* **5**, eaav7574 (2019).
60. Baykov, A. A. Energy coupling in cation-pumping pyrophosphatase—Back to Mitchell. *Front Plant Sci* **11**, 107 (2020).
61. Baykov, A. A., Anashkin, V. A., Malinen, A. M. & Bogachev, A. V. The mechanism of energy coupling in H^+/Na^+-pumping membrane pyrophosphatase—possibilities and probabilities. *Int J Mol Sci* **23**, 9504 (2022).
62. Anashkin, V. A. *et al*. Rapid kinetics of H^+ transport by membrane pyrophosphatase: evidence for a "direct-coupling" mechanism. *Biochem Biophys Res Commun* **744**, 151203 (2025).

63. Bogachev, A. V., Anashkin, V. A., Bertsova, Y. V., Zavyalova, E. G. & Baykov, A. A. Na^+ translocation dominates over H^+-translocation in the membrane pyrophosphatase with dual transport specificity. *Int J Mol Sci* **25**, 11963 (2024).
64. Ryan, S. J., Carlson, C. J., Mordecai, E. A. & Johnson, L. R. Global expansion and redistribution of Aedes-borne virus transmission risk with climate change. *PLoS Negl Trop Dis* **13**, e0007213 (2019).
65. World Health Organization. *World health statistics 2023: monitoring health for the SDGs, sustainable development goals* (World Health Organization, 2023).
66. Zhang, Y. *et al*. Structural isosteres of phosphate groups in the protein data bank. *J Chem Inf Model* **57**, 499–516 (2017).
67. Johansson, N. G. *et al*. Discovery of membrane-bound pyrophosphatase inhibitors derived from an isoxazole fragment. *ACS Med Chem Lett* **11**, 605–610 (2020).
68. Johansson, N. G. *et al*. Exploration of Pyrazolo [1, 5-a] pyrimidines as membrane-bound pyrophosphatase inhibitors. *ChemMedChem* **16**, 3360–3367 (2021).
69. Jeschke, G. The contribution of modern EPR to structural biology. *Emerg Top Life Sci* **2**, 9–18 (2018).
70. Schiemann, O. & Prisner, T. F. Long-range distance determinations in biomacromolecules by EPR spectroscopy. *Q Rev Biophys* **40**, 1–53 (2007).
71. Goldfarb, D. Exploring protein conformations in vitro and in cell with EPR distance measurements. *Curr Opin Struct Biol* **75**, 102398 (2022).
72. Kapsalis, C. *et al*. Allosteric activation of an ion channel triggered by modification of mechanosensitive nano-pockets. *Nat Commun* **10**, 4619 (2019).
73. Kapsalis, C., Ma, Y., Bode, B. E. & Pliotas, C. In-lipid structure of pressure-sensitive domains hints mechanosensitive channel functional diversity. *Biophys J* **119**, 448–459 (2020).
74. Gopinath, A., Rath, T., Morgner, N. & Joseph, B. Lateral gating mechanism and plasticity of the β-barrel assembly machinery complex in micelles and *Escherichia coli*. *PNAS Nexus* **3**, 19 (2024).
75. Galazzo, L. *et al*. The ABC transporter MsbA adopts the wide inward-open conformation in *E. coli* cells. *Sci Adv* **8**, eabn6845 (2022).
76. Thaker, T. M. *et al*. Asymmetric drug binding in an ATP-loaded inward-facing state of an ABC transporter. *Nat Chem Biol* **18**, 226–235 (2022).
77. Wingler, L. M. *et al*. Angiotensin analogs with divergent bias stabilize distinct receptor conformations. *Cell* **176**, 468–478 (2019).
78. Haysom, S. F. *et al*. Darobactin B stabilises a lateral-closed conformation of the BAM complex in *E. coli* cells. *Angew Chem Int Ed Engl* **62**, e202218783 (2023).
79. McHaourab, H. S., Steed, P. R. & Kazmier, K. Toward the fourth dimension of membrane protein structure: insight into dynamics from spin-labeling EPR spectroscopy. *Structure* **19**, 1549–1561 (2011).
80. Bordignon, E., Kucher, S. & Polyhach, Y. EPR techniques to probe insertion and conformation of spin-labeled proteins in lipid bilayers in *Lipid-protein interactions: methods and protocols* (ed. Kleinschmidt, J. H.) 493–528 (Springer New York, 2019).
81. Hartley, A. M., Ma, Y., Lane, B. J., Wang, B. & Pliotas, C. Using pulsed EPR in the structural analysis of integral membrane proteins in *Electron paramagnetic resonance: Volume 27* (eds. Chechik, V., Murphy, D. M. & Bode, B. E.) 74–108 (The Royal Society of Chemistry, 2020).
82. Pliotas, C. *et al*. Conformational state of the MscS mechanosensitive channel in solution revealed by pulsed electron-electron double resonance (PELDOR) spectroscopy. *Proc Natl Acad Sci USA* **109**, E2675–E2682 (2012).
83. Hagelueken, G., Ward, R., Naismith, J. H. & Schiemann, O. MtsslWizard: in silico spin-labeling and generation of distance distributions in PyMOL. *Appl Magn Reson* **42**, 377–391 (2012).
84. Branigan, E., Pliotas, C., Hagelueken, G. & Naismith, J. H. Quantification of free cysteines in membrane and soluble proteins using a fluorescent dye and thermal unfolding. *Nat Protoc* **8**, 2090–2097 (2013).

85. Pliotas, C. Ion channel conformation and oligomerization assessment by site-directed spin labeling and pulsed-EPR. *Method Enzymol* **594**, 203–242 (2017).
86. Lane, B. J. *et al*. HDX-guided EPR spectroscopy to interrogate membrane protein dynamics. *STAR Protoc* **3**, 101562 (2022).
87. Peter, M. F. *et al*. Cross-validation of distance measurements in proteins by PELDOR/DEER and single-molecule FRET. *Nat Commun* **13**, 4396 (2022).
88. Jeschke, G. *et al*. DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Appl Magn Reson* **30**, 473–498 (2006).
89. Chiang, Y. W., Borbat, P. P. & Freed, J. H. The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J Magn Reson* **172**, 279–295 (2005).
90. Liebschner, D. Protein crystallography-methods and protocols. *Struct Biol Commun* **74**, 74–75 (2018).
91. Ishchenko, A., Abola, E. & Cherezov, V. Lipidic cubic phase technologies for structural studies of membrane proteins in *Membrane proteins production for structural analysis* (ed. Mus-Veteau, I.) 289–314 (Springer New York, 2014).
92. Landau, E. M. & Rosenbusch, J. P. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci USA* **93**, 14532–14535 (1996).
93. Bragg, W. H. The reflection of X-rays by crystals.(II.). *Proc R Soc Lond A* **89**, 246–248 (1913).
94. Drenth, J. The theory of X-ray diffraction by a crystal in *Principles of protein X-ray crystallography* (ed. Drenth, J.) 70–116 (Springer New York, 1999).
95. Taylor, G. L. Introduction to phasing. *Biol Crystallogr* **66**, 325–338 (2010).
96. Guss, J. M. *et al*. Phase determination by multiple-wavelength x-ray diffraction: crystal structure of a basic "blue" copper protein from cucumbers. *Science* **241**, 806–811 (1988).
97. de La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol* **276**, 472–494 (1997).
98. Rose, J. P. & Wang, B. C. SAD phasing: history, current impact and future opportunities. *Arch Biochem Biophys* **602**, 80–94 (2016).
99. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Biol Crystallogr* **66**, 486–501 (2010).
100. Adams, P. D. *et al*. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Biol Crystallogr* **66**, 213–221 (2010).
101. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26**, 283–291 (1993).
102. Srinivasan, B. Words of advice: teaching macromolecular crystallography. *FEBS J* **290**, 5441–5455 (2023).
103. Kühlbrandt, W. The resolution revolution. *Science* **343**, 1443–1444 (2014).
104. Glaeser, R. M. How good can single-particle cryo-EM become? What remains before it approaches its physical limits?. *Annu Rev Biophys* **48**, 45–61 (2019).
105. Dubochet, J. *et al*. Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys* **21**, 129–228 (1988).
106. Tivol, W. F., Briegel, A. & Jensen, G. J. An improved cryogen for plunge freezing. *Microsc Microanal* **14**, 375–379 (2008).
107. Haynes, R. M. *et al*. A strategic approach for efficient cryo-EM grid optimization using design of experiments. *J Struct Biol* **217**, 108068 (2025).
108. Kampjut, D., Steiner, J. & Sazanov, L. A. Cryo-EM grid optimization for membrane proteins. *iScience* **24**, 102139 (2021).
109. Frank, J. & Al-Ali, L. Signal-to-noise ratio of electron micrographs obtained by cross correlation. *Nature* **256**, 376–379 (1975).
110. Bendory, T., Bartesaghi, A. & Singer, A. Single-particle cryo-electron microscopy: Mathematical theory, computational challenges, and opportunities. *IEEE Signal Process Mag* **37**, 58–76 (2020).

111. Frank, J. Single-particle reconstruction of biological molecules-story in a sample. *Angew Chem Int Ed Engl* 57, 10826–10841 (2018).
112. Wu, X. & Rapoport, T. A. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proc Natl Acad Sci USA* **118**, e2115001118 (2021).
113. Zheng, S. Q. *et al*. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331–332 (2017).
114. Sheth, L. K., Piotrowski, A. L. & Voss, N. R. Visualization and quality assessment of the contrast transfer function estimation. *J Struct Biol* **192**, 222–234 (2015).
115. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519–530 (2012).
116. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296 (2017).
117. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216–221 (2015).
118. Bazzone, A. & Barthmes, M. Functional characterization of SLC transporters using solid supported membranes in *Biophysics of membrane proteins: methods and protocols* (eds. Vincent, L. G. P. & Goldman, A) 73–103 (Humana, 2020).
119. Bazzone, A., Barthmes, M. & Fendler, K. SSM-based electrophysiology for transporter research. *Methods Enzymol* **594**, 31–83 (2017).
120. Bazzone, A. *et al*. A comparative study on the lysosomal cation channel TMEM175 using automated whole-cell patch-clamp, lysosomal patch-clamp, and solid supported membrane-based electrophysiology: functional characterization and high-throughput screening assay development. *Int J Mol Sci* **24**, 12788 (2023).
121. Dondapati, S. K. *et al*. Functional reconstitution of membrane proteins derived from eukaryotic cell-free systems. *Front Pharmacol* **10**, 917 (2019).
122. Mavridou, V., King, M. S., Bazzone, A., Springett, R. & Kunji, E. R. Membrane potential stimulates ADP import and ATP export by the mitochondrial ADP/ATP carrier due to its positively charged binding site. *Sci Adv* **10**, eadp7725 (2024).
123. Kornreich, B. G. The patch clamp technique: principles and technical considerations. *J Vet Cardiol* **9**, 25–37 (2007).
124. Liu, T. H. *et al*. The proximity between C-termini of dimeric vacuolar H^+-pyrophosphatase determined using atomic force microscopy and a gold nanoparticle technique. *FEBS J* **276**, 4381–4394 (2009).
125. Bazzone, A., Barthmes, M. & Fendler, K. SSM-based electrophysiology for transporter research. *Methods Enzymol* **594**, 31–83 (2017).
126. McCoy, A. J. *et al*. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).
127. Russell, H., Cura, R. & Lovett, J. E. DEER data analysis software: a comparative guide. *Front Mol Biosci* **9**, 915167 (2022).
128. Schiemann, O. *et al*. Benchmark test and guidelines for DEER/PELDOR experiments on nitroxide-labeled biomolecules. *J Am Chem Soc* **143**, 17875–17890 (2021).
129. Development Core R Team. *R: a language and environment for statistical computing* (Massachusetts Institute of Technology, 2014).
130. Commo, F. & Bot, B. M. R package nplr n-parameter logistic regressions. *CRAN*. https://cran.uvigo.es/web/packages/nplr/vignettes/nplr.pdf (2016).
131. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671–675 (2012).
132. Mukherjee, S. *et al*. Synthetic antibodies against BRIL as universal fiducial marks for single-particle cryoEM structure determination of membrane proteins. *Nat Commun* **11**, 1598 (2020).
133. de la Rosa-Trevin, J. M. *et al*. Scipion: a software framework toward integration, reproducibility and validation in 3D electron microscopy. *J Struct Biol* **195**, 93–99 (2016).
134. Klose, D. *et al*. Resolving distance variations by single-molecule FRET and EPR spectroscopy using rotamer libraries. *Biophys J* **120**, 4842–4858 (2021).

135. Bhattacharyya, A. On a measure of divergence between two multinomial populations. *Sankhyā* **7**, 401–406 (1946).
136. McIntosh, M. T., Drozdowicz, Y. M., Laroiya, K., Rea, P. A. & Vaidya, A. B. Two classes of plant-like vacuolar-type H^+-pyrophosphatases in malaria parasites. *Mol Biochem Parasitol* **114**, 183–195 (2001).
137. Lee, S. *et al*. How do branched detergents stabilize GPCRs in micelles?. *Biochemistry* **59**, 2125–2134 (2020).
138. Stretton, S., Techkarnjanaruk, S., Mclennan, A. M. & Goodman, A. E. Use of green fluorescent protein to tag and investigate gene expression in marine bacteria. *Appl Environ Microbiol* **64**, 2554–2559 (1998).
139. Gordon-Weeks, R., Parmar, S., Davies, T. E. & Leigh, R. A. Structural aspects of the effectiveness of bisphosphonates as competitive inhibitors of the plant vacuolar proton-pumping pyrophosphatase. *Biochem J* **337**, 373–377 (1999).
140. Nadezhdin, K. D. *et al*. Structural mechanisms of TRPM7 activation and inhibition. *Nat Commun* **14**, 2639 (2023).
141. Li, S. Detergents and alternatives in cryo-EM studies of membrane proteins: Detergents and alternatives in cryo-EM studies of membrane proteins. *Acta Biochim Biophys Sin (Shanghai)* **54**, 1049–1056 (2022).
142. Lee, H. J., Lee, H. S., Youn, T., Byrne, B. & Chae, P. S. Impact of novel detergents on membrane protein studies. *Chem* **8**, 980–1013 (2022).
143. Nikolova, M. S. *Mechanism and structure of the PACE family of transport proteins* (Ph.D thesis, University of Leeds, 2022).
144. Zhang, K., Wu, H., Hoppe, N., Manglik, A. & Cheng, Y. Fusion protein strategies for cryo-EM study of G protein-coupled receptors. *Nat Commun* **13**, 4366 (2022).
145. Nygaard, R., Kim, J. & Mancia, F. Cryo-electron microscopy analysis of small membrane proteins. *Curr Opin Struct Biol* **64**, 26–33 (2020).
146. Kim, J. *et al*. Structure and drug resistance of the *Plasmodium falciparum* transporter PfCRT. *Nature* **576**, 315–320 (2019).
147. Hirst, I. J., Thomas, W. J., Davies, R. A. & Muench, S. P. CryoEM grid preparation: a closer look at advancements and impact of preparation mode and new approaches. *Biochem Soc Trans* **52**, 1529–1537 (2024).
148. Noble, A. J. *et al*. Reducing effects of particle adsorption to the air-water interface in cryo-EM. *Nat Methods* **15**, 793–795 (2018).
149. Efremov, R. G., Leitner, A., Aebersold, R. & Raunser, S. Architecture and conformational switch mechanism of the ryanodine receptor. *Nature* **517**, 39–43 (2015).
150. Wang, L. *et al*. Structure and mechanogating of the mammalian tactile channel PIEZO2. *Nature* **573**, 225–229 (2019).
151. Denisov, I. G. & Sligar, S. G. Nanodiscs for structural and functional studies of membrane proteins. *Nat Struct Mol Biol* **23**, 481–486 (2016).
152. Hesketh, S. J. *et al*. Styrene maleic-acid lipid particles (SMALPs) into detergent or amphipols: an exchange protocol for membrane protein characterisation. *Biochim Biophys Acta Biomembr* **1862**, 183192 (2020).
153. Breyton, C. *et al*. Assemblies of lauryl maltose neopentyl glycol (LMNG) and LMNG-solubilized membrane proteins. *Biochim Biophys Acta Biomembr* **1861**, 939–957 (2019).
154. Rasmussen, S. G. *et al*. Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature* **469**, 175–180 (2011).